

RESEARCH

Open Access



Technical efficiency of provincial public healthcare in South Africa

Victor Ngobeni^{*} , Marthinus C. Breitenbach and Goodness C. Aye

Abstract

Background: Forty-nine million people or 83 per cent of the entire population of 59 million rely on the public healthcare system in South Africa. Coupled with a shortage of medical professionals, high migration, inequality and unemployment; healthcare provision is under extreme pressure. Due to negligence by the health professionals, provincial health departments had medical-legal claims estimated at R80 billion in 2017/18. In the same period, provincial health spending accounted for 33 per cent of total provincial expenditure of R570.3 billion or 6 per cent of South Africa's Gross Domestic Product. Despite this, healthcare outcomes are poor and provinces are inefficient in the use of the allocated funds. This warrants a scientific investigation into the technical efficiency of the public health system.

Methods: The study uses data envelopment analysis (DEA) to assess the technical efficiency of the nine South African provinces in the provision of healthcare. This is achieved by determining, assessing and comparing ways that individual provinces can benchmark their performance against peers to improve efficiency scores. DEA compares firms operating in homogenous conditions in the usage of multiple inputs to produce multiple outputs. Therefore, DEA is ideal for measuring the technical efficiency of provinces in the provision of public healthcare. In DEA methodology, the firms with scores of 100 per cent are technically efficient and those with scores lower than 100 per cent are technically inefficient. This study considers six DEA models using the 2017/18 total health spending and health staff as inputs and the infant mortality rate as an output. The first three models assume the constant returns to scale (CRS) while the last three use the variable return to scale (VRS) both with an input-minimisation objective.

Results: The study found the mean technical efficiency scores ranging from 35.7 to 87.2 per cent between the health models 1 and 6. Therefore, inefficient provinces could improve the use of inputs within a range of 64.3 and 20.8 per cent. The Gauteng province defines the technical efficiency frontiers in all the six models. The second-best performing province is the North West province. Other provinces like KwaZulu-Natal, Limpopo and the Eastern Cape only perform well under the VRS. The other three provinces are inefficient.

Conclusions: Based on the VRS models 4 to 6, the study presents three policy options. Policy option 1 (model 4): the efficiency gains from addressing health expenditure wastage in four inefficient provinces amounts to R17 billion. Policy option 2 (model 5): the potential savings from the same provinces could be obtained from reducing 17,000 health personnel, advisably, in non-core areas. In terms of Policy option 3 (model 6), three inefficient provinces should reduce 6940 health workers while the same provinces, inclusive of KwaZulu-Natal could realise health expenditure savings of R61 million. The potential resource savings from improving the efficiency of the inefficient provinces could be used to refurbish and build more hospitals to alleviate pressure on the public health system. This could also reduce the per capita numbers per public hospital and perhaps their performance as overcrowding is reportedly negatively affecting their performance and health outcomes. The potential savings could also be used to appoint and train medical

*Correspondence: victorngobeni@me.com

Department of Economics, University of Pretoria, Lynnwood Rd, Hatfield, Pretoria 0002, South Africa



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

practitioners, specialists and researchers to reduce the alarming numbers of medical legal claims. Given the existing challenges, South Africa is not ready to implement the National Health Insurance (NHI) Scheme, as it requires additional financial and human resources. Instead, huge improvements in public healthcare provision could be achieved by re-allocating the resources 'saved' through efficiency measures by increasing the quality of public healthcare and extending healthcare to more recipients.

Keywords: Expenditure, Data envelopment analysis, Healthcare, Inefficiency, Technical efficiency

JEL Classification: C6-Mathematical programming models, D2-Production and organisations, Fiscal policies and behaviour of economic agents, I1-Health

Introduction

According to Statistics South Africa [57], South Africa is a Southern African country with a population of 55.6 million, estimated at 59 million in 2018. Coovadia et al. [15] and Mayosi and Benatar [39] indicate that South Africa is a middle-income country with health outcomes worse than in many low-income countries. This is exacerbated by inadequate human resources to cater for a growing population coupled with a rising number of refugees and economic migrants. As a result, there is growing concern about the state of the public health system, its efficiency and capability to provide sustainable services. The South African health care sector is comprised of public and private segments that are deep-rooted in the past unjust policies of Apartheid, which caused disparities in healthcare access between black and white citizens that there is inequitable access to health services between the poor majority and the rich minority in South Africa. This situation still persists. However, the current divides are generally between the rich and the poor irrespective of race.

The Government of the Republic of South Africa [23] states that, overall, the health sector has 813 hospitals with 133,387 beds for acute health care. The public sector accounts for 49.7 per cent or 404 of these hospitals with 69 per cent of total bed allocation. The private sector comprises 409 hospitals (50.3 per cent) with 31 per cent of the total bed allocation. These numbers clearly point to inequality between the private and public hospitals and may further hint at a shortage of public health infrastructure as these numbers translate into 81,188 people per public hospital with an average of 228 beds per public facility. Mayosi and Benatar [39] add that many state hospitals are in a dire state with much of public healthcare infrastructure run down and dysfunctional due to underfunding, mismanagement, and neglect. This compromises the quality of healthcare and leads to earlier than required patient discharges. The Competition Commission [14] states that, in 2018, the vulnerable and poorly resourced public healthcare facilities served approximately 83 per cent of the population who were largely without medical insurance. The private healthcare facilities served 17 per cent of the population private

healthcare insurance. According to the Government of the Republic of South Africa [23], there is a substantial difference in resource availability between the public and private health sectors with more than half of financial and human resources allocated to the private sector.

Marten et al. [36] indicate that, aside from the institutional structure of the healthcare system and inadequate public health infrastructure, another Achilles heel is the absolute and chronic deficit of healthcare workers in the country. Table 1 shows that the existing healthcare workers are unevenly distributed along the health qualification categories and geographical areas. This uneven distribution of staff and skills according to Coovadia et al. [15] has compromised the ability to deliver key programmes, notably for Human Immunodeficiency Virus (HIV), tuberculosis (TB), child health, mental health, and maternal health. Mayosi and Benatar [39] mention that nurses are central to healthcare, especially in rural areas where physicians are reluctant to practice. Table 1 also shows that the nursing personnel accounts for 143,264 or 63 per cent of total health personnel while medical practitioners, specialists and researchers account for 19,988 or 9 per cent. Despite a small component of medical doctors as a proportion of total healthcare workers, 70 per cent are employed in the private sector, implying shortages in the public sector. Moreover, government's increased investment in the medical profession produced more doctors over the years, but a brain drain has since reversed these gains. South Africa incurs the highest costs for medical doctor education but, in turn, loses returns on investment as doctors migrate to Europe. Thirty per cent of South African doctors have emigrated and 58 per cent are intending to emigrate to Western countries. Community health workers account for 23.7 per cent or 54,180 of health personnel and 4.3 per cent to other health personnel categories.

Another major challenge in the healthcare sector is the high cost of providing medical care in South Africa. The Health Systems Trust [25] indicates that this cost is already high and has been increasing rapidly over the last three decades. Moreover, the high public sector salaries and excessive administration costs; duplication of

Table 1 Practicing health personnel by sector April 2018. Source: Author's own Table based on Health Systems Trust [24]

Province	Pupil auxiliary nurses	Student nurses	Enrolled nurses	Nursing assistants	Professional nurses	Medical practitioners	Medical researchers	Medical specialists	Other	Total	% of total
Eastern Cape	356	–	3263	5260	10,993	1903	1	177	6210	28,163	12
Free State	76	–	939	1972	2295	664	5	293	2954	9198	4
Gauteng	823	2902	7694	6518	14,223	3614	13	1929	12,034	49,750	22
KwaZulu-Natal	512	951	9926	5976	17,163	3383	120	808	12,480	51,319	22
Limpopo	64	470	4085	4731	9409	1248	12	60	12,389	32,468	14
Mpumalanga	67	749	1881	1431	5471	1079	1	78	7687	18,444	8
Northern Cape	116	–	237	864	1520	457	2	21	3070	6287	3
North West	49	21	958	2489	4511	934	–	116	6842	15,920	7
Western Cape	225	–	2608	4152	5314	1719	6	1345	5622	20,991	9
Total	2288	5093	31,591	33,393	70,899	15,001	160	4827	64,965	228,217	100

"Other" refers to clinical associates, community health workers, dental practitioners, therapists and specialists, environmental health workers, occupational therapists, pharmacists, physiotherapists, psychologists, radiographers and speech therapists and audiologists. The figures exclude national personnel. Community health workers comprise 83 per cent of the other category

services and inefficiencies are a serious problem. Table 2 shows that the provincial health sector accounted for R186.9 billion of the total provincial budget in 2017/18, which was a 6 per cent contribution to the country's GDP by all provinces for the same period. The total health spending of R186.9 billion was also equivalent to 33 per cent of total provincial expenditure of R570.3 billion in the same year. The compensation of employees' budget accounted for 61 per cent of total health expenditure.

Coovadia et al. [15] report that the health sector is also negatively affected by weak political and management leadership. There is insufficient political and management leadership to manage underperformance, especially in the public health sector. This negatively impacts on the efficient administration of the sector and on the provision of quality healthcare services. These challenges have to be resolved urgently, as the Government of the Republic of South Africa [23] reports that South Africa is currently working towards the provision of free quality universal healthcare (UHC) by 2030 which will be publicly funded. Therefore, UHC could require additional financial and human resources prompting the efficient use of existing funding resources. Given the substantial health budgets and their impact on macroeconomic indicators and human development, Verhoeven et al. [60] maintain that there is a general concern about the rapid rise in costs and the trade-offs between efficiency and equity. The Health Systems Trust [25] states that the vision of UHC and the current inefficient delivery of health services are mutually exclusive. As a result, cost containment is an important consideration in the delivery of healthcare as financial and human resources available for healthcare are limited, especially in the public

sector where competition for resources is greater, creating a gap between available and required resources for healthcare delivery. Moreover, the National Treasury [47] reports that provincial health departments also face contingent liability risks associated with medical-legal claims due to negligence by health professionals. In 2017/18, this liability was estimated at R80 billion, up 32 per cent from 2016/17. Pay-outs against these claims amounted to R1.5 billion in 2017/18 and are projected to exceed R2 billion in 2018/19.

The Health Systems Trust [25] states that, while it is important to improve the efficiency of existing resources, there is limited available information for efficiency determination. Moreover, scientific economic analysis is not usually used when allocating resources. Given the limited funding of the public healthcare sector and increasing healthcare expenditure requirements, the use of scientific methods to evaluate and compare public healthcare efficiency spending is inevitable and critical in reshaping healthcare policy. The present study fills this gap by using DEA to determine the technical efficiency of provincial healthcare in South Africa. Coelli et al. [13] state that DEA is widely used to measure technical efficiency in the public sector, including in the healthcare sector. Despite this, studies measuring technical efficiency of provincial healthcare sector in South Africa are lacking. Most international studies capturing South Africa compare the national health system with that of other countries without outlining specific provincial health efficiency or inefficiency levels. This is despite the provincial health spending accounting for 33 per cent of provincial annual budgets. A survey of the health sector DEA efficiency assessment literature that was conducted in the current study reflects that the study is the first to assess and quantify in-depth, the technical efficiency of the provincial healthcare sector in South Africa. Therefore, the study generates new knowledge for efficiency benchmarking and improvement. Given the background of the health sector, the rest of the paper is set out as follows: in "Literature review" section, the study conducts a review of literature using DEA to assess healthcare efficiency in the public sector, "Methodology and data" section outlines methodology and data, "Efficiency results" section discusses the efficiency results and "Conclusion" section summarises the findings of the study and details its recommendations.

Literature review

In terms of DEA literature, in Europe, DEA was used by Campanella et al. [11] to assess the technical efficiency of 50 Italian hospitals. The results revealed an average efficiency score of 77 per cent amongst the DMUs, with a requirement for the inefficient DMUs to reduce

Table 2 Provincial health contributions to Gross Domestic Product. Sources: Statistics South Africa [56]. National Treasury [42–46, 48–51]

Province	GDP	Health spending	Health spending % GDP
Eastern Cape	247,040,000	22,771,139	9
Free State	154,400,000	9,795,191	6
Gauteng	1,080,800,000	44,132,368	4
KwaZulu-Natal	494,080,000	40,430,163	8
Limpopo	216,160,000	19,522,743	9
Mpumalanga	216,160,000	12,445,693	6
Northern Cape	61,760,000	4,722,157	8
North West	185,280,000	11,420,212	6
Western Cape	432,320,000	21,671,137	5
Total	3,088,000,000	186,910,803	6

GDP Gross Domestic Product in R'000 in 2017 terms

their input usage by 23 per cent to achieve efficiency. Lo Storto and Goncharuk [34] employed DEA to measure the technical efficiency of 32 European (EU) countries. The results for model 1 showed that inefficient countries should reduce input usage by 36 and 34 per cent in 2011 and 2014, respectively. The mean efficiency score for model 2 indicated that inefficient DMUs should increase outputs by 68 and 56 per cent for the same years. Asandului et al. [6] also used DEA to analyse the efficiency of healthcare systems of 30 EU countries. The model 1 results respectively showed an average efficiency score of 74 and 75 per cent under the assumptions of CRS and VRS. The model 2 results yielded slightly different average technical efficiency scores of 81 and 77 per cent respectively for the CRS and VRS perspectives. Asandului and Fatulescu [7] used DEA to evaluate the technical efficiency of the healthcare systems of 27 EU countries. The study identified five countries on the efficiency frontier. 14 countries had efficiency scores below 50 per cent, needing to reduce their input usage by more than 50 per cent to be efficient. The remaining countries had efficiency levels of above 50 per cent. Another technical efficiency analytical study in the EU region was conducted by Baciú and Bolezat [9] who assessed the technical efficiency of healthcare of 27 EU countries. The average technical efficiency score of all the DMUs was 60 per cent, showing more room by the inefficient DMUs to reduce the overall input usage by 40 per cent while maintaining the same output levels. Anton [4] measured the technical efficiency of 20 healthcare systems in Eastern and Central Europe. The study obtained an overall mean efficiency score of 98 per cent for life expectancy and 82 per cent for the infant mortality of all the DMUs in the studied healthcare systems. Varabyova and Schreyögg [59] assessed the technical efficiency of healthcare in 31 Organisation for Economic Co-operation and Development (OECD) countries using DEA. The mean efficiency score was 70 per cent, indicating possible input contraction by 30 per cent and possible output expansions by the same proportions. Afonso and St. Aubyn [1] applied DEA to compare the output efficiency of healthcare systems of 30 OECD countries. They ascertained that seven countries were on average technically efficient and inefficient countries could on average increase their output efficiency by 40 per cent. Chowdhury et al. [12] analysed the technical efficiency of 113 acute healthcare hospitals in Canada, a North American country, revealing that most Ontario hospitals were not technically efficient, 65 per cent were subject to decreasing returns to scale (DRS), 10 per cent to increasing return to scale (IRS) and 25 per cent were CRS efficient. Gannon [20] assessed the technical efficiency of 60 Irish hospitals using DEA. The results showed a mean efficiency score of 96 per cent,

thereby requiring inefficient DMUs to curb input wastage by 4 per cent.

Other international DEA studies that involved Africa and, in particular, South Africa, are also presented in this section. Alhassan et al. [3] used DEA to assess the technical efficiency of 64 health facilities in Ghana. The results showed an average technical efficiency score of 65 per cent, meaning a reduction in the use of inputs by 35 per cent. DEA was also used by Jarjue et al. [26] to determine the technical efficiency of 41 secondary healthcare centres in the Gambia. The mean efficiency score was 65 per cent, meaning that inefficient DMUs could still increase output by 35 per cent. 10 per cent of the DMUs were scale efficient while 90 per cent were scale inefficient with a mean efficiency score of 87 per cent. Kim and Kang [27] applied DEA to analyse the technical efficiency of healthcare systems of 170 countries, including South Africa. The research found that only 17 per cent of the studied countries used inputs efficiently. Asian countries were the most efficient and South Africa was amongst the inefficient countries, with an efficiency score of 84 per cent wasting about 16 per cent of inputs. High and upper middle-income countries had efficiency scores of over 70 per cent and lower-middle income and lower-income countries recorded the average efficiency scores of 67 per cent and 66 per cent respectively. Additional research by Prasetyo and Zuhdi [54] investigated the technical efficiency of healthcare provision in 81 countries, including South Africa. The study observed that 17 countries were efficient in using government expenditure and South Africa had an efficiency score of 94 per cent. DEA was also used by Marschall and Flessa [35] to compute the technical efficiency of 20 healthcare centres in Burkina Faso. The CRS approach found an average mean efficiency score of 91 per cent and the VRS approach yielded a mean efficiency score of 94 per cent; reflecting a potential increase in outputs by 9 and 6 per cent respectively. Akazili et al. [2] applied DEA to calculate the technical efficiency of 89 randomly sampled healthcare centres in Ghana. They found that 65 per cent of healthcare centres were technically inefficient as they used the resources they did not need. They had an average technical efficiency score of 57 per cent, implying that on average, they could reduce input utilisation by 43 per cent without reducing the prevailing output levels.

In another case, DEA was adopted by Benneyan et al. [10] to compare the technical efficiency of healthcare systems of 180 countries. They found that 115 countries, including South Africa were not efficient. Masiye [38] used DEA to evaluate the technical efficiency of 30 hospitals in Zambia. The overall results showed that Zambian hospitals operated at a 67 per cent level of technical efficiency, implying that 33 per cent of input resources were

being wasted. Zere et al. [65] adopted an input-minimisation objective to analyse the technical efficiency of 30 district hospitals in Namibia. The CRS results yielded the average efficiency scores ranging between 63 and 74 per cent over the study period. The mean VRS efficiency scores ranged from 67 to 72 per cent. The study also revealed that IRS were a predominant form of scale inefficiency. Kirigia et al. [28] analysed the technical efficiency of 155 public clinics in KwaZulu-Natal in South Africa. They ascertained that only 47 clinics were technically efficient and 108 were not. The inefficient clinics had room to be efficient by reducing inputs and increasing outputs. They could reduce the number of nurses by 417 and general staff by 457. These reductions represented 31 and 32 per cent inefficiency rates. The outputs could be increased by 115,534 antenatal care visits, 1010 baby deliveries, 179,075 child health care visits, 5702 dental care visits, 121,658 family planning visits, 56,068 psychiatry visits, 34,270 sexually transmitted infections related care visits and 34,270 TB related visits. Lastly, Kirigia et al. [29] assessed the technical efficiency of 55 provincial hospitals in KwaZulu-Natal using DEA. The average technical efficiency score was 90.6 per cent. 22 DMUs were inefficient needing to reduce inputs as follows, 117 doctors, 295 administration staff, 2709 nurses, 835 general staff, 1191 provisioning staff, 61 paramedics, 58 technicians, 38 other staff members and 1752 beds.

In terms of the 21 health sector studies reviewed in this paper, only five related to South Africa. Three of the studies were comparative cross-country efficiency benchmarking studies than relative provincial efficiency analytical studies. These studies used different variables to those applied in the present study. The other two studies compared the efficiency of clinics and hospitals in one of the nine South African province, however, they did not use expenditure as a variable of analysis. The current study focuses on the efficiency of all the provinces, it is specific and explicit in terms of which provinces should be prioritised for reforms, therefore, enabling benchmarking and improvement.

Methodology and data

In this paper, we follow the DEA approach developed by Farrell in 1957 and enhanced in 1978 by Charnes, Cooper and Rhodes (also called the CCR model) to convert the fractional linear efficiency estimates into linear mathematical efficiency programmes under the CRS. We also use the VRS approach reported by Gavurova et al. [22] to have been developed in 1984 by Banker, Charnes and Cooper to transform the CCR model to allow for consideration of scale efficiency analysis. This is called the Banker, Charnes and Cooper (BCC) model. The terminology “envelopment” in DEA refers to the ability of the efficiency production

frontier to tightly enclose the production technology (input and output variables). Cooper et al. [16] and McWilliams et al. [40] state that DEA was developed in a microeconomic setting and applied to firms to convert inputs into outputs. However, in efficiency determination, the term “firm” is often replaced by the more encompassing DMU. DEA is an appropriate method of computing and analysing the efficiency of public sector institutions as they employ multiple inputs to produce multiple outputs. DEA estimated frontiers do not make any assumptions upfront or specify the functional form related to the production technology. Aristovnik [5] and Martić et al. [37] state that there are input-minimisation and output-maximisation DEA orientation models. The former determines the quantity of inputs that could be curtailed without reducing the prevailing level of outputs to make the DMUs efficient and the latter expands outputs of the DMUs until a combination of inputs and outputs reach the production possibility frontier while holding the levels of inputs constant. However, the selection of each orientation is dependent on the objectives of a particular study.

According to Taylor and Harris [58], DEA is a comparative efficiency measurement tool that evaluates the efficiency of homogeneous DMUs operating in similar environmental conditions and where there is no known relationship between the conversion of inputs and outputs. Wang and Alvi [61] report that DEA only uses the information used in a particular study to determine efficiency and does not take into consideration other factors that are exogenous to the study. DEA measures the distance or derivatives of production functions to determine the extent of DMU's efficiency deviation from the optimal position. It classifies the DMUs into extremely efficient performers versus inefficient performers. In terms of the DEA methodology, the current study uses both the CCR and the BCC models to test for stability, variability and robustness of the obtained efficiency results. These models are described in the following paragraphs.

Under the CCR model, suppose there are M different number of inputs and P different number of outputs for N DMUs. These quantities are represented by column vectors x_{ij} ($i = 1, 2, 3, \dots, M, j = 1, 2, 3, \dots, N$) and q_{rj} ($r = 1, 2, 3, \dots, P, j = 1, 2, 3, \dots, N$). The $M \times N$ input matrix, X and $P \times N$ output matrix, Q represents the production technology for all the N number of DMUs. For each DMU, the ratio of all the output variables over all the input variables is represented by $u'q_{rj}/v'x_i$. Where $u = P \times 1$ vector output weights and $v = M \times 1$ vector input weights. The optimal weights or the efficiency estimates are obtained by solving a mathematical problem. In the context of the CRS, an efficient DMU operates at most productive scale size (MPSS) or technically optimal production scale (TOPS). Hence, the optimal weights or

efficiency estimates are obtained by solving a mathematical problem that is reflected in Eq. 1.

$$\begin{aligned} \text{Top}_s &= \max_{u,v} (u'q_{rj}/v'x_{ij}) \\ \text{St.} \\ u'q_{rj}/v'x_{ij} &\leq 1 \\ u, v &\geq 0 \end{aligned} \quad (1)$$

Equation 1 shows the original linear programme, called the primal. It aims to maximise the efficiency score, which is represented by the ratio of all the weights of outputs to inputs, subject to the efficiency score not exceeding 1, with all inputs and outputs being positive. Equation 1, has an infinite number of solutions, if (u, v) is a solution, so is αv . To avoid this, one can impose a constraint $v'x_{ij} = 1$, which produces Eq. 2.

$$\begin{aligned} \max_{u,v} (u'q_{rj}) \\ \text{St.} \\ v'x_{ij} &= 1 \\ u'q_{rj} - v'x_{ij} &\leq 0 \\ u, v &\geq 0 \end{aligned} \quad (2)$$

An equivalent envelopment problem can be developed for the problem in Eq. 2, using duality in linear programming. The dual for $\max_{u,v} (u'q_{rj})$ is $\min \theta, \lambda \theta$. The value of θ is the efficiency score; it satisfies the condition $\theta \leq 1$; it is the scalar measure. Lauro et al. [32] report that λ is an $N \times 1$ vector of all constants representing intensity variables indicating necessary combinations of efficient entities or reference units (peers) for every inefficient DMU, it limits the efficiency of each DMU to be greater than 1. This results in Eq. 3, which represents the CCR-CRS model with an input minimisation orientation.

$$\begin{aligned} \min \theta, \lambda \theta \\ \text{St.} \\ -q_{rj} + Q\lambda &\geq 0 \\ \theta x_i - X\lambda &\geq 0 \\ \lambda &\geq 0. \end{aligned} \quad (3)$$

Avkiran [8] states that the CRS postulates no significant relationship between DMU's operational size and their efficiency. That is, under the CRS assumption, the large DMUs are deemed to attain the same levels of efficiency as small DMUs in transforming inputs to outputs. Therefore, the CRS assumption implies that the size of a DMU is not relevant when assessing technical efficiency. However, in most cases, DMUs have varying sizes and this becomes a factor when determining their efficiency. As a result, Gavurova et al. [22] mention that in 1984, the CCR formulation was generalised to allow for the VRS. Lavado and Domingo [31] argued that, in practice, there is plethora of factors such as financial constraints that may result in the DMUs not operating at optimal scale. Aristovnik [5] adds that, if one cannot assume the existence of the CRS, then a VRS type of DEA is an appropriate choice for computing efficiency. Gannon [20] advises that the VRS should be used if it is likely that the size of a DMU will have a bearing on efficiency. As such, Yawe [63] cautions that the use of the CRS specification when the DMUs are not operating at an optimal scale results in a measure of technical efficiency which is confounded by scale effects. The solution is to use the VRS as it permits for the calculation of scale inefficiency. The CRS linear programming problem can be modified to account for the VRS by adding the convexity constraint: $N1'\lambda = 1$ to Eq. 3, where $N1$ is an $N \times 1$ vector of ones to formulate Eq. 4. Equation 4 represents the BBC-VRS model with an input-minimisation orientation. Therefore, Eqs. 1 to 3 represent the CRS models while Eqs. 4 to 5 represents the VRS models.

$$\begin{aligned} \min \theta, \lambda \theta \\ \text{St.} \\ -q_{rj} + Q\lambda &\geq 0 \\ \theta x_{ij} - X\lambda &\geq 0 \\ N1'\lambda &= 1 \\ \lambda &\geq 0 \end{aligned} \quad (4)$$

Lauro et al. [32] and Yuan and Shan [64] report that the CCR and the BCC models only differ in the manner the latter includes convexity constraints. Since the current model considers the VRS, the restriction $\sum_{i=1}^n \lambda_i = 1$ is introduced. Ramírez Hassan [55] cautions that, if this restriction is not there, it would imply the application of the CRS model. The same analogy applies to all the inefficient provinces in the sample. That is, the slacks and the radial movements are calculated for all inefficient provinces using Eq. 5. The

BCC is adept to calculate pure technical efficiency and inefficiency and when applied with the CCR model, it also measures scale inefficiency. Where, $\sum_{i=1}^I \lambda_i = 1$, a DMU is on a CRS frontier, if $\sum_{i=1}^I \lambda_i < 1$, the DMU is located on the IRS frontier and if $\sum_{i=1}^I \lambda_i > 1$, there is DRS. Given that this study has adopted both the CCR and the VRS with an input-minimisation orientation. The DEA models used in this study also consider the slack movements for the inefficient DMUs. As a result, the models account for the slacks in Eq. 5.

$$\begin{aligned} & \text{Min} \theta, \lambda_j, S_r^+, S_i^- \\ & \theta - \varepsilon \left[\sum_{i=1}^M S_i^- + \sum_{r=1}^P S_r^+ \right] \\ & \theta x_{i0} - \sum_{j=1}^N x_{ij} \lambda_j - S_i^- = 0, \quad \sum_{j=1}^N x_{ij} \lambda_j = X \lambda \\ & \theta q_{r0} = \sum_{j=1}^N q_{rj} \lambda_j - S_r^+ = 0, \quad \sum_{j=1}^N q_{rj} \lambda_j = Q \lambda \\ & \sum_{j=1}^N \lambda_j = 1 \\ & \lambda_j, S_r^+, S_i^- > 0 \end{aligned} \quad (5)$$

Coelli et al. [13] define slacks as input excesses and output shortfalls that are required over and above the initial radial movements to push DMUs to efficiency levels. Both the slack and radial movements are characterised only with the inefficient DMUs. The radial movements are initial input contractions or output expansions that are required for a firm to become efficient. S_i^+ and S_i^- in Eq. 5 are the output and input slacks respectively to be calculated with θ , and λ_j . ε , is the non-Archimedean constant. Gavurova et al. [22] hint that if the slack variables of a DMU are not equal to zero and the technical efficiency score is lower than one, it is necessary to perform a non-radial shift that is expressed by the slack variables to achieve technical efficiency. In Eq. 5, the slack variables determine the optimum level of inputs that DMUs would have to utilise and the outputs that they would have to produce to become efficient, provided that these DMUs are inefficient. Therefore, the slacks depict the under-produced outputs or overused inputs.

Given the specified model, it is clear that this paper uses the CCR and BCC models with input-minimisation objectives to analyse technical efficiency of provincial healthcare in South Africa for the 2017/18 financial year. This is

relevant for this study as the technical efficiency of provincial healthcare considers total health spending and total health staff as inputs. The DMUs have control over these variables, especially on expenditure as opposed to health outputs. The selected output variable for this study is the IMR. A longitudinal approach was not adopted since the composition of provincial health expenditure remains the same throughout. Table 3 summarises the efficiency analytical data. The literature review presented in this paper shows that total health spending, health staff and the IMRs are commonly used variables to analyse health sector efficiency. Therefore, this paper considers various health production technologies with a maximum three-variables. Data for the IMR and total health staff are actual figures obtained from the audited annual reports of provinces for 2017/18. Data for total health spending is obtained from the National Treasury's 2017/18 provincial revenue and expenditure data. There are high variations reflected by high standard deviations from the mean with respect to health expenditure across all the provinces. The minimum expenditure of R4.7 billion was recorded in the Northern Cape, with Gauteng spending the sample maximum of

Table 3 Input and output variables. Sources: Eastern Cape Department of Health [17], Free State Department of Health [18], Gauteng Department of Health [21], KwaZulu-Natal Department of Health [30], Limpopo Department of Health [33], Mpumalanga Department of Health [41], Northern Cape Department of Health [52], North West Department of Health [53], Western Cape Government Health [62]. National Treasury [42–46, 48–51]

Provinces	Health inputs		Health output
	x1 (THE)	x2 (THS)	y1 (IMR)
Eastern Cape	22,771,139	40,424	14
Free State	9,795,191	17,301	11.8
Gauteng	44,132,368	66,124	10.1
KwaZulu-Natal	40,430,163	68,125	12.4
Limpopo	19,522,743	33,848	12.4
Mpumalanga	12,445,693	20,421	9.7
Northern Cape	4,722,157	6924	11.6
North West	11,420,212	17,536	8.1
Western Cape	21,671,137	31,549	9.3
Observations	9	9	9
Mean	20,767,867	33,584	11
Minimum	4,722,157	6924	8
Maximum	44,132,368	68,125	14
Median	19,522,743	31,549	12
Standard deviation	12,794,907	20,302	2

THE is the actual total health expenditure or spending measured in R'000

THS is the total number of people or workers employed in the health sector

IMR refers to the number of deaths per 1000 live births of children under 1 year of age

R44.1 billion. There were also major variations in the number of total health staff per province.

The six health models considered in this study are summarised in Table 4, of which three are based on the CRS and three on the VRS. The shapes of the frontiers of these models are illustrated graphically by Figs. 1, 2, 3, 4, 5 and 6 in Appendix.

Efficiency results

Avrikan [8] reports that the CRS efficiency scores represent technical efficiency. On the other hand, the VRS efficiency scores represent pure technical efficiency. Fried et al. [19] provide intuition that the efficiency results calculated through the VRS are always higher than those calculated using the CRS. This is because the best VRS efficiency frontier is only formed by convex efficient combinations of inputs and outputs. As a result, the BCC model envelops data tighter than the CCR model. Moreover, the VRS is comprehensive as it captures DMUs that are efficient in the CRS and VRS dispensations. Therefore, it contains a smaller number of inefficient DMUs. Table 5 shows the provincial health technical and scale efficiency scores for all the six health models. The health model 1 yielded an average health technical efficiency score of 35.7 per cent, implying that all the inefficient provinces should reduce total health spending by 64.3 per cent while maintaining the same levels of the IMR. Gauteng defined the health efficiency frontier and the North West province was very close to optimality with an efficiency score of 80 per cent. The other seven DMUs had the efficiency scores ranging between 6.4 and 30 per cent, implying inefficiency range of between 70 and 93.6 per cent.

The Eastern Cape was the least efficient province with an efficiency score of 6.4 per cent. Mpumalanga records an efficiency score of 12.9 per cent, the Free State 13.7 per cent, the Northern Cape 18.3 per cent, KwaZulu-Natal, Limpopo and the Western Cape had efficiency scores of 30 per cent each. As it relates to health model 2, the mean efficiency score was 35.4 per cent. The efficiency scores for all the DMUs, except for the Eastern Cape were similar to those in health model 1. The Eastern Cape was still

the worst performing DMU with an efficiency score of 3.5 per cent. Table 5 reveals an average efficiency score of 63.9 per cent for health model 3 when total health expenditure and total health staff were used together as inputs of the health production frontier while maintaining the same rate of IMR. These variables complemented each other; the use of one without the other decreased the efficiency scores. In this model, Gauteng, the Northern Cape and the North West Provinces were technically efficient. The Free State, Limpopo, Mpumalanga and the Western Cape's efficiency scores surpassed 50 per cent. KwaZulu-Natal and the Eastern Cape were below 50 per cent with the latter as an extreme inefficient outlier.

When only total health spending was used as an input under the VRS in health model 4, the mean efficiency score was 69 per cent. This was 33.3 per cent higher than the efficiency results that are obtained in health model 1 for the same variable. This implied that the size of provinces matters in determining the technical efficiency of their health spending. Five provinces, the Eastern Cape, Gauteng, KwaZulu-Natal, Limpopo and the North West Provinces were purely technically efficient. This means that four provinces had to reduce total health expenditure inefficiency by 31 per cent. This model also showed that the Eastern Cape, KwaZulu-Natal, North West and Limpopo were largely disadvantaged when scale was disregarded, as they were only efficient under the VRS while they were inefficient under the CRS. The Mpumalanga Province was the least efficient DMU in this scenario with an inefficiency score of 85.7 per cent. It was followed by the Free State, Western Cape and the Northern Cape Provinces with inefficiency rates of 68.7, 66.7 and 58.3 per cent respectively.

The frontier for total health staff in health model 5 was exactly similar to the one for total health spending in health model 4, meaning that the efficiency scores for the DMUs were similar when individually assessing these two variables. As a result, the shapes of the efficiency frontiers of health models 4 and 5 were also similar. This implied that either one of these variables was appropriate to assess the technical efficiency of the health sector under

Table 4 Health efficiency DEA models

Models	DEA model	Number of variables	Variable description
Health model 1	CRS	2	Total health expenditure and infant mortality rate
Health model 1	CRS	2	Total health staff and infant mortality rate
Health model 1	CRS	3	Total health expenditure, total health staff and infant mortality rate
Health model 1	VRS	2	Total health expenditure and infant mortality rate
Health model 1	VRS	2	Total health staff and infant mortality rate
Health model 1	VRS	3	Total health expenditure, total health staff and infant mortality rate

CRS constant returns to scale and VRS = Variable returns to scale

Table 5 Health technical and scale efficiency scores. Source: Author's graph based on DEAP 2.1 efficiency results

Province	Health model 1	Health model 2	Health model 3	Health model 4	Health model 5	Health model 6	Scale efficiency					
Eastern Cape	0.064	0.035	0.063	1000	1000	1000	0.064	DRS	0.035	DRS	0.063	DRS
Free State	0.137	0.137	0.607	0.313	0.313	0.670	0.440	DRS	0.440	DRS	0.906	DRS
Gauteng	1.000	1000	1000	1000	1000	1000	1000	–	1000	–	1000	–
KwaZulu-Natal	0.300	0.300	0.375	1000	1000	1000	0.300	DRS	0.300	DRS	0.375	DRS
Limpopo	0.300	0.300	0.648	1000	1000	1000	0.300	DRS	0.300	DRS	0.648	DRS
Mpumalanga	0.129	0.129	0.504	0.143	0.143	0.542	0.900	IRS	0.900	IRS	0.930	IRS
Northern Cape	0.183	0.183	1000	0.417	0.417	1.000	0.440	DRS	0.440	DRS	1.000	–
North West	0.800	0.800	1000	1000	1000	1000	0.800	IRS	0.800	IRS	1.000	–
Western Cape	0.300	0.300	0.552	0.333	0.333	0.635	0.900	IRS	0.900	IRS	0.870	DRS
Mean	0.357	0.354	0.639	0.690	0.690	0.872	0.572		0.568		0.755	
CRS constant returns to scale, VRS variable return to scale												

CRS constant returns to scale, VRS variable return to scale

the VRS. The results of the health model 6 also show that when both these input variables were considered under the VRS, the average pure technical efficiency score of the DMUs increased by 23.3 per cent compared to health model 3 results. The health model 6 mean efficiency score was 87.2 per cent, denoting that inefficient DMUs should reduce total health expenditure and total health personnel by 12.8 per cent. Six provinces were purely technically efficient in health model 6. These were the Eastern Cape, Gauteng, KwaZulu-Natal, Limpopo, Northern Cape and the North-West. The inefficient DMUs are the Free State, Mpumalanga and the Western Cape.

Table 5 also shows the scale efficiency scores for the DMUs under consideration. The first column under the heading “scale efficiency” in Table 5 shows the scale efficiency scores derived by dividing the efficiency results of health model 1 by health model 4 results. The second column shows the type of scale efficiency. Therefore, the average scale efficiency score of all the DMUs when total health spending was used as a single variable was 57.2 per cent, depicting high levels of scale inefficiency. The scale inefficient DMUs need to improve scale efficiency by 42.8 per cent. Only Gauteng was CRS scale efficient, providing a benchmark for all the scale inefficient DMUs. The Western Cape, Mpumalanga and the North West Provinces were very close to scale efficiency. They were the only DMUs on the IRS frontier while the other five scale inefficient DMUs were on a DRS curve. The third column in Table 5 indicates an average scale inefficiency score of 56.8 per cent when total health personnel was used as a single input, resulting in 8 DMUs being scale inefficient. The Western Cape, Mpumalanga and the North West Provinces are very close to scale efficiency. They were the only DMUs on IRS frontier while the other five scale inefficient DMUs were on DRS. The last two columns in Table 5 show that when both inputs were employed, the average scale efficiency was 75.5 per cent. Gauteng, Northern Cape and the North West were scale efficient, Mpumalanga was the only DMU on the IRS while the other scale inefficient provinces were on DRS.

Table 6 indicates the radial and slack movements for all the health DEA models. In the health model 1, the mean efficiency score of 35.7 per cent translated into an average inefficiency score of 64.3 per cent. This inefficiency score was equivalent to eight inefficient provinces overusing total health expenditure by R46.4 billion. In other words, total health spending could be reduced by R46.4 billion, while still producing the same output levels. The Eastern Cape could reduce total health spending by R20.6 billion (93.6 per cent inefficiency rate) given the IMR of 14 per cent. The Free State’s and Mpumalanga’s respective inefficient total health spending levels were R6.9 billion (86.3 per cent inefficiency score) and R6.1 billion (87.1 per cent rate of inefficiency) given their current levels of IMR of 11 and 9

per cent. The Northern Cape had overall health spending radial movement of R4.9 billion (81.7 per cent inefficiency score), accounting for 10.6 per cent of the consolidated provincial health spending inefficiency with an IMR of 11 per cent. KwaZulu-Natal, Limpopo, Western Cape and the North West Provinces collectively accounted for 17 per cent or R7.9 billion (associated with their respective 70, 70, 70 and 80 per cent inefficiency rates) of the overall health expenditure inefficiency at the prevailing IMR. In health model 2, all the provinces had a mean health staff inefficiency score of 64.6 per cent, equivalent to 64,400 more health personnel than required. The Eastern Cape had to reduce health workers by 38,600 (96.5 inefficiency weight) while still maintaining the IMR at 14 per cent. The Free State had 6900 (86.3 per cent inefficiency weight) more personnel than it should, Mpumalanga 6100 (87.1 per cent inefficiency score) and the Northern Cape 4900 (81.7 per cent inefficiency score) at the prevailing IMRs. KwaZulu-Natal and Limpopo were each supposed to reduce their total health workers by 2800 (30 per cent efficiency rate) and the Western Cape by 2100 (30 per cent efficiency score). The North West was closest to the efficiency frontier defined by Gauteng, it only had 200 (20 per cent inefficiency rate) more health personnel than required. In terms of health model 3, when total health spending and total health staff were simultaneously applied as inputs, the overall health expenditure mean inefficiency score of 36.1 per cent was equal to R481.8 million total health spending and 49,337 total health personnel in line with an improvement in the efficiency scores from 35 per cent to 63.9 per cent. The Eastern Cape had R397.2 million overall health expenditure inefficiency faced with a daunting task of reducing health sector employees by 37,470 (93.7 per cent inefficiency rate). KwaZulu-Natal was required to curtail total health spending by 42.5 million and total health staff by 2500 in line with a 62.5 per cent inefficiency score to become efficient. All inefficient provinces had to reduce the average technical inefficiency score of 36.1 per cent for both health spending and staff. The Free State used R6.7 million more than it should and could reduce the overall health personnel by 3148 to address its inefficiency rate of 39.3 per cent. The Limpopo province had health expenditure and staff radial movements of R11.6 million and 1407 respectively, associated with its inefficiency rate of 35.2 per cent. The Mpumalanga Province had an inefficiency score of 49.6 per cent; it overused total health spending of R9.9 million and should reduce total health staff by 3469 to become efficient. The Western Cape should reduce health expenditure by R13.9 million and health staff by 1343 (44.8 per cent inefficiency score). The peers for the Eastern Cape, Free State, Limpopo, Mpumalanga and the Western Cape are the North West and the Northern Cape and for KwaZulu-Natal is Gauteng.

Table 6 Total health expenditure and health staff radial and slack movements: CRS. Source: Author's graph based on DEAP 2.1 efficiency results

Provinces	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	Total
Model 1: CRS THE only										
Original input	22,000,000	8,000,000	1,000,000	4,000,000	4,000,000	7,000,000	6,000,000	1,000,000	3,000,000	56,000,000
Input radial movement	(20,600,000)	(6,900,000)	–	(2,800,000)	(2,800,000)	(6,100,000)	(4,900,000)	(2000)	(2,100,000)	(46,400,000)
Input slack movement	–	–	–	–	–	–	–	–	–	–
Input target	1,400,000	1,100,000	1,000,000	1,200,000	1,200,000	900,000	1,100,000	800,000	900,000	9,600,000
Original output	14	11	10	12	12	9	11	8	9	96
Output radial movement	–	–	–	–	–	–	–	–	–	–
Output target	14	11	10	12	12	9	11	8	9	96
DMU peers	3	3	3	3	3	3	3	3	3	3
Model 2: CRS THS only										
Original input	40,000	8000	1000	4000	4000	7000	6000	1000	3000	74,000
Input radial movement	(38,600)	(6900)	–	(2800)	(2800)	(6100)	(4900)	(200)	(2100)	(64,400)
Input slack movement	–	–	–	–	–	–	–	–	–	–
Input target	1400	1100	1000	1200	1200	900	1100	800	900	9600
Original output	14	11	10	12	12	9	11	8	9	96
Output radial movement	–	–	–	–	–	–	–	–	–	–
Output target	14	11	10	12	12	9	11	8	9	96
DMU peers	3	3	3	3	3	3	3	3	3	3
Model 3: CRS THS and THE										
THS original input	40,000	8000	1000	4000	4000	7000	6000	1000	3000	74,000
THS radial movement	(37,470)	(3148)	–	(2500)	(1407)	(3469)	–	–	(1343)	49,337
THS slack movement	–	–	–	–	–	–	–	–	–	–
THS target	2530	4852	1000	1500	2593	3531	6000	1000	1657	24,663
THE original input	424,000	17,000	66,000	68,000	33,000	20,000	6,000	17,000	31,000	682,000
THE radial movement	(397,181)	(6689)	–	(42,500)	(11,607)	(9912)	–	–	(13,875)	481,764
THE slack movement	–	–	–	–	–	–	–	–	–	–
THE target	26,819	10,311	66,000	25,500	21,393	10,088	6000	17,000	17,125	200,236
Original output	14	11	10	12	12	9	11	8	9	96
Output radial movement	–	–	–	–	–	–	–	–	–	–
Output target	14	11	10	12	12	9	11	8	9	96
DMU peers	7;8	7;8	3	8	7;8	7;8	7	8	7;8	7;8

THE is in R'000

DMU decision-making unit, THE total health expenditure, THS total health staff, CRS constant return to scale

Table 7 shows that in terms of the health model 4, only five DMUs were purely technically inefficient with an overall inefficiency score of 31 per cent, translating into total health spending inefficiency of R17 billion. Mpumalanga accounted for R6 billion (85.7 per cent inefficiency score) of this amount, the Free State for R5.5 billion (68.7 per cent inefficiency weight), Northern Cape for R3.5 billion (58.3 per cent inefficiency score) and the Western Cape for R2 billion (66.7 per cent inefficiency score). This overspending could be curtailed while maintaining the same IMRs. The Free State and the Northern Cape should draw lessons from Gauteng and KwaZulu-Natal while Mpumalanga, North West and the Western Cape should learn from Gauteng. The model showed output slacks of 1, 2 and 1 per cent for Mpumalanga, North West and the Western Cape respectively. Given the nature of the IMR, it is not ideal to increase this measure. The same five provinces were inefficient in health model 5. Their mean inefficiency score of 31 per cent was equivalent to 17,000 more health workers than required. The Free State accounted for 32.4 per cent or 5500 of this amount (68.7 per cent inefficiency score), Mpumalanga 35.3 per cent or 6000 (85.7 per cent inefficiency score), Northern Cape 20.6 per cent or 3500 (58.3 per cent) and the Western Cape for 11.8 per cent or 2000 (66.7 per cent inefficiency score) of excess health staff. The model showed the output slacks of 1, 2 and 1 per cent for Mpumalanga, North West and the Western Cape respectively. The Free State's peers were Gauteng and KwaZulu-Natal and the rest of the DMUs could draw lessons from Gauteng. In terms of health model 6, only three provinces had a mean inefficiency score of 12.8 per cent, equivalent to inefficient total health spending of R26.1 million and 6940 excess health staff. The Northern Cape reached the efficiency point. The Free State had excess staff of 2644 and R5.6 million of health expenditure (33 per cent inefficiency rate), Mpumalanga 3202 and R9.2 million (45.8 per cent inefficiency score) and the Western Cape of 1094 and R11.3 million (36.5 per cent inefficiency weight). The model showed an output slack of 0.7 per cent for Mpumalanga. The Free State and the Western Cape could improve their performance by benchmarking with Limpopo, Northern Cape and North West while Mpumalanga's peers are the Northern Cape and the North West.

Conclusion

This study determined the average technical efficiency scores of the six health models. The mean efficiency scores ranged from 35.7 per cent to 87.2 per cent between health model 1 and health model 6. This illustrates that, in line with theoretical postulates, the DEA results under the VRS were higher than under the CRS. The sensitivity analysis was conducted by dropping one input variable in the first two models both under the CRS and the VRS

and reintroducing it in the last models under the same assumptions. This exercise showed that the efficiency results improved when the two variables were used simultaneously and decreased when they were used individually under the CRS and the VRS. This further reflects that the health efficiency results are sensitive to the number of inputs used in efficiency analysis. Using a single health input generated lower efficiency results in both assumptions. Figure 7 in [Appendix](#) illustrates the individual provincial scores within the broader DMUs' performance. Gauteng was the best performing DMU defining the efficiency frontier in all the six health models. The second-best performing province was the North West. Other provinces like KwaZulu-Natal, Limpopo and the Eastern Cape only perform well under the VRS. Mpumalanga, Western Cape and the Free State are poorly performing provinces with respect to health sector efficiency.

The efficiency scores of the VRS model (models 4, 5 and 6) are used to formulate the recommendations for this study since this method is comprehensive. The following policy options and recommendations are made.

- Policy option 1 (health model 4): Target minimising total health expenditure in four inefficient provinces, the Free State, Mpumalanga, Northern Cape and the Western Cape. Their collective healthcare expenditure inefficiencies amounted to R17 billion. The Free State should curtail overspending by R5.5 billion and the Northern Cape's is R3.5 billion. They could draw lessons from Gauteng and KwaZulu-Natal. Mpumalanga was spending R6 billion on healthcare more than it should while the Western Cape should reduce its inefficiency by R2 billion. All the interventions to realise these levels of spending should be implemented while maintaining the prevailing levels of IMRs. These two provinces can benchmark their health operations with Gauteng for pure technical efficiency improvements.
- Policy option 2 (health model 5): Target reducing the overusage of the health staff in the four inefficient provinces. In terms of minimising total health staff, there was overstaffing in the Free State, Mpumalanga, Northern Cape and the Western Cape. Their consolidated health staff inefficiency was 17,000 people. The Free State should reduce health staff complement by 5500 benchmarking with Gauteng and KwaZulu-Natal, Mpumalanga by 7000, Northern Cape by 6000 and the Western Cape by 2000 with all of them benchmarking with Gauteng while maintaining the same rates infant mortality.
- Policy option 3 (health model 6): In cases where the interest of policy makers was to improve health staff and expenditure at the same time, the Free State,

Table 7 Total health expenditure and health staff radial and slack movements: VRS. Source: DEAP 2.1 efficiency results

Provinces	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	Total
Model 4: VRS THE only										
Original input	22,000,000	8,000,000	1,000,000	4,000,000	4,000,000	7,000,000	6,000,000	1,000,000	3,000,000	56,000,000
Input radial movement	–	(5,500,000)	–	–	–	(6,000,000)	(3,500,000)	–	(2,000,000)	(17,000,000)
Input slack movement	–	–	–	–	–	–	–	–	–	–
Input target	22,000,000	2,500,000	1,000,000	4,000,000	4,000,000	1,000,000	2,500,000	1,000,000	1,000,000	39,000,000
Original output	14	11	10	12	12	9	11	8	9	96
Output radial movement	–	–	–	–	–	–	–	–	–	–
Output slack movement	–	–	–	–	–	1	–	2	1	4
Output target	14	11	10	12	12	10	11	10	10	10
DMU peers	1	4;3	3	4	5	3	4;3	3	3	3
Model 5: VRS THS only										
Original input	40,000	8000	1000	4000	4000	7000	6000	1000	3000	74,000
Input radial movement	–	(5500)	–	–	–	(6000)	(3500)	–	(2000)	(17,000)
Input slack movement	–	–	–	–	–	–	–	–	–	–
Input target	40,000	2500	1000	4000	4000	1000	2500	1000	1000	57,000
Original output	14	11	10	12	12	9	11	8	9	96
Output radial movement	–	–	–	–	–	–	–	–	–	–
Output slack movement	–	–	–	–	–	1	–	2	1	4
Output target	14	11	10	12	12	10	11	10	10	100
DMU peers	1	3;4	3	4	4	3	3	3	3	3
Model 6: VRS THS and THE										
THS original input	40,000	8000	1000	4000	4000	7000	6000	1000	3000	74,000
THS radial movement	–	(2644)	–	–	–	(3202)	–	–	(1094)	(6940)
THS slack movement	–	–	–	–	–	–	–	–	–	–
THS target	40,000	5356	1000	4000	4000	3798	6000	1000	1906	67,060
THE original input	424,000	17,000	66,000	68,000	33,000	20,000	6000	17,000	31,000	682,000
THE radial movement	–	(5618)	–	–	–	(9153)	–	–	(11,305)	(26,076)
THE slack movement	–	–	–	(35,000)	–	–	–	–	–	(35,000)
THE target	424,000	11,382	66,000	33,000	33,000	10,847	6000	17,000	19,695	620,924
Original output	14	11	10	12	12	9	11	8	9	96
Output radial movement	–	–	–	–	–	–	–	–	–	–
Output slack movement	–	–	–	–	–	0.678	–	–	–	0.678
Output target	14	11	10	12	12	9.7	11	8	9	97
DMU peers	1	5;8;7	3	5	5	7;8	7	8	7;5;8	7

DMU decision-making unit, THE total health expenditure, THS total health staff, VRS variable returns to scale

Mpumalanga and the Western Cape should respectively reduce staff by 2644, 3202 and 1094 while maintaining the same levels of IMR. The same provinces should be assisted to cut their spending inefficiencies by R26.1 million. The Free State by R5.6 million, Mpumalanga by R9.2 million and the Western Cape by R11.3 million while KwaZulu-Natal should curtail THE by R35 million. The Free State and the Western Cape should learn from Limpopo, the Northern Cape and the North West to improve their efficiency while Mpumalanga should benchmark with the last two provinces.

The potential savings from improving the efficiency of the inefficient provinces could also be used to refurbish and build more hospitals to alleviate the pressure on the public health system. This could also reduce the per capita numbers per public hospital and perhaps their performance as overcrowding is reportedly negatively affecting their performance and health outcomes. Moreover, overcrowded hospitals amid low professional health workers place pressure on the few appointed core health staff complement. This warrants the use of potential savings to appoint more personnel, especially medical practitioners, specialists and researchers while reducing personnel expenditure in non-core areas, as there is a general shortage of healthcare practitioners in South Africa. This implies the improvement of the ratio of practitioners and nursing assistants to attended patients. Moreover, it is essential to retrain health professionals using the realised efficiency gains to reduce the alarming numbers of medical-legal claims. This could also free up additional resources to enhance service delivery.

The study also cautions that, given healthcare personnel and infrastructure challenges, South Africa is not ready to implement the National Health Insurance (NHI) scheme. The scheme requires additional financial and human resources amidst the existing inefficiencies. Instead of taking on the NHI at this juncture, which would be very costly, South Africa can make huge improvements in public healthcare provision by improving efficiency and re-allocating those resources 'saved' through efficiency measures, to improving the quality of healthcare and extending healthcare to more recipients. Implementing the NHI without implementing the efficiency measures will set up the health sector for certain failure. In general, provinces should also review the high spending levels on goods and services to ensure value-for-money.

Inefficient provinces should continuously monitor the efficiency of health spending and health personnel and publish their efficiency performance periodically for public scrutiny. This could promote efficient and evidence-based budgeting. In terms of study limitations, this study

only determines the inefficiency levels of health spending and health personnel without methodologically explaining the factors resulting in these inefficiencies. As a result, it is difficult to understand why such inefficiencies exist. The study also suffers from a small sample size problem by just analysing nine provinces. However, this limitation is structural given that South Africa has only nine provinces.

Acknowledgements

I hereby acknowledge the assistance and guidance of Professor Marthinus Breitenbach and Dr. Goodness Aye for their contributions to the study.

Authors' contributions

VN is the main author as a Doctorate students under the supervision of MB and GA. MB and GA suggested the appropriate analytical model while VN inputted the data and interpreted the results. MB and GA checked the accuracy of the model results, their interpretation and also contributed on the recommendations. All authors read and approved the final manuscript.

Funding

The authors hereby acknowledges the financial assistance of the National Research Foundation and the University of Pretoria. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily attributed to the National Research Foundation.

Availability of data and materials

The datasets analysed during the current study are available from the corresponding author on reasonable request. This particularly relates to the DEA model inputs and results. Most of data sources are listed in the reference list.

Ethics approval and consent to participate

Not applicable.

Consent for publication

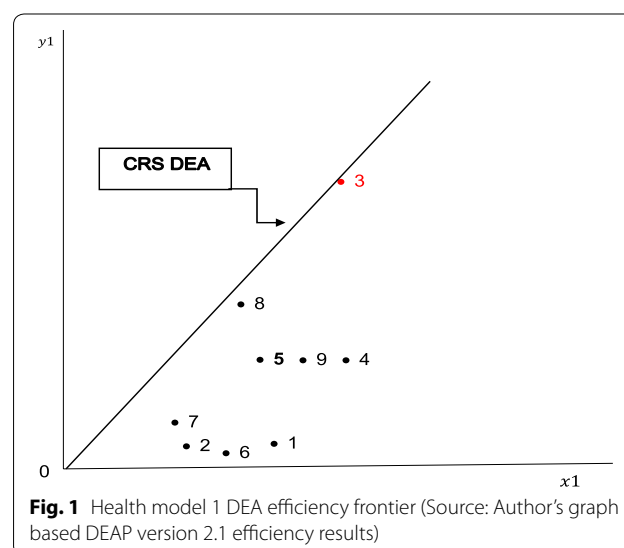
Not applicable.

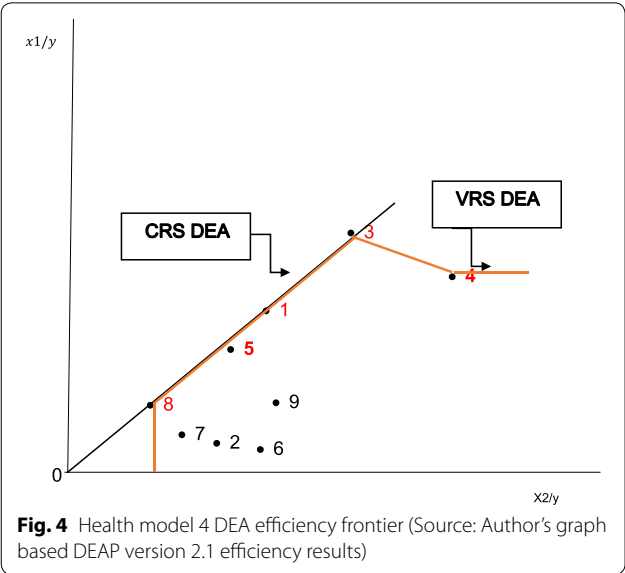
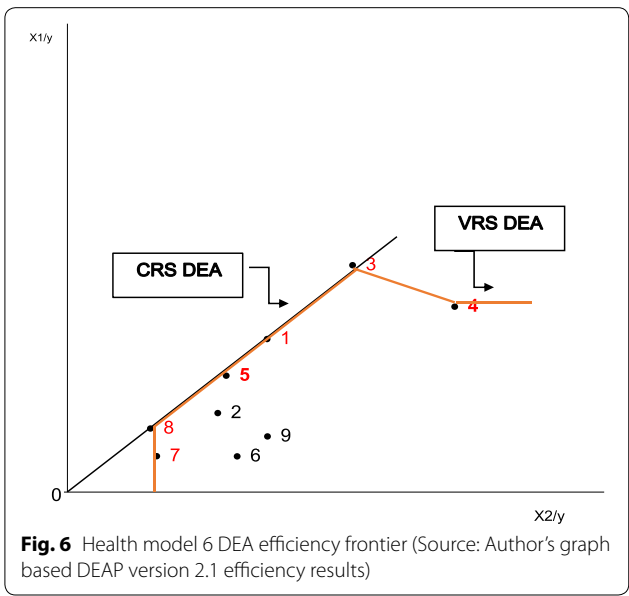
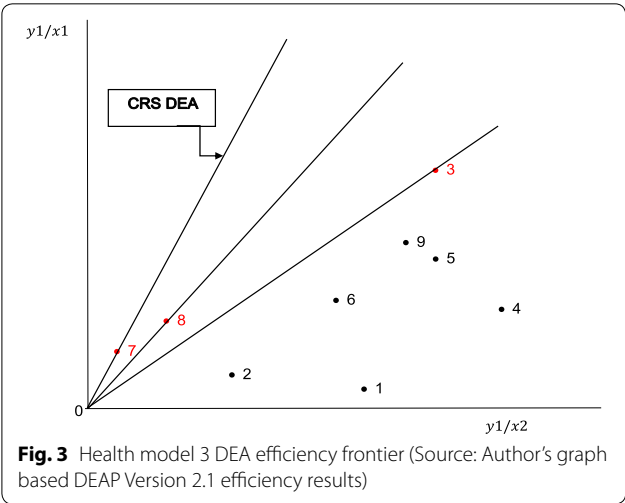
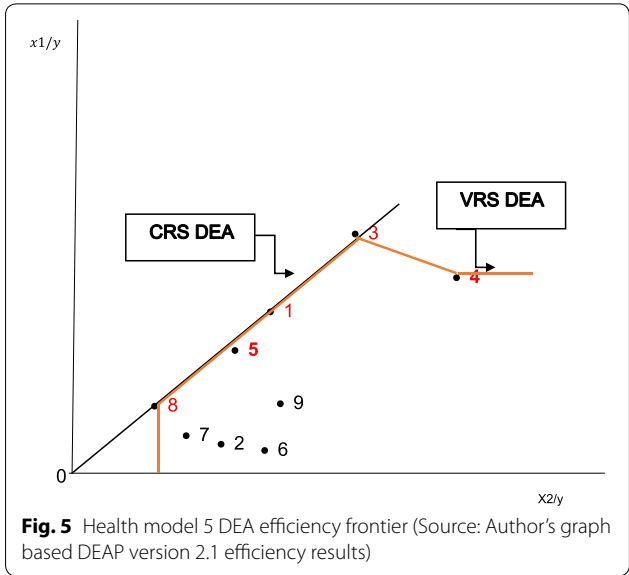
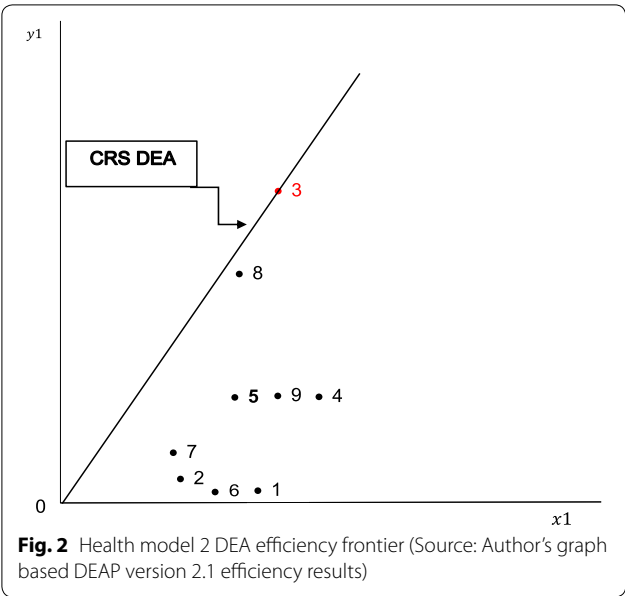
Competing interests

The authors declare that they have no competing interests.

Appendix: Health efficiency frontiers

See Figs. 1, 2, 3, 4, 5, 6, 7.





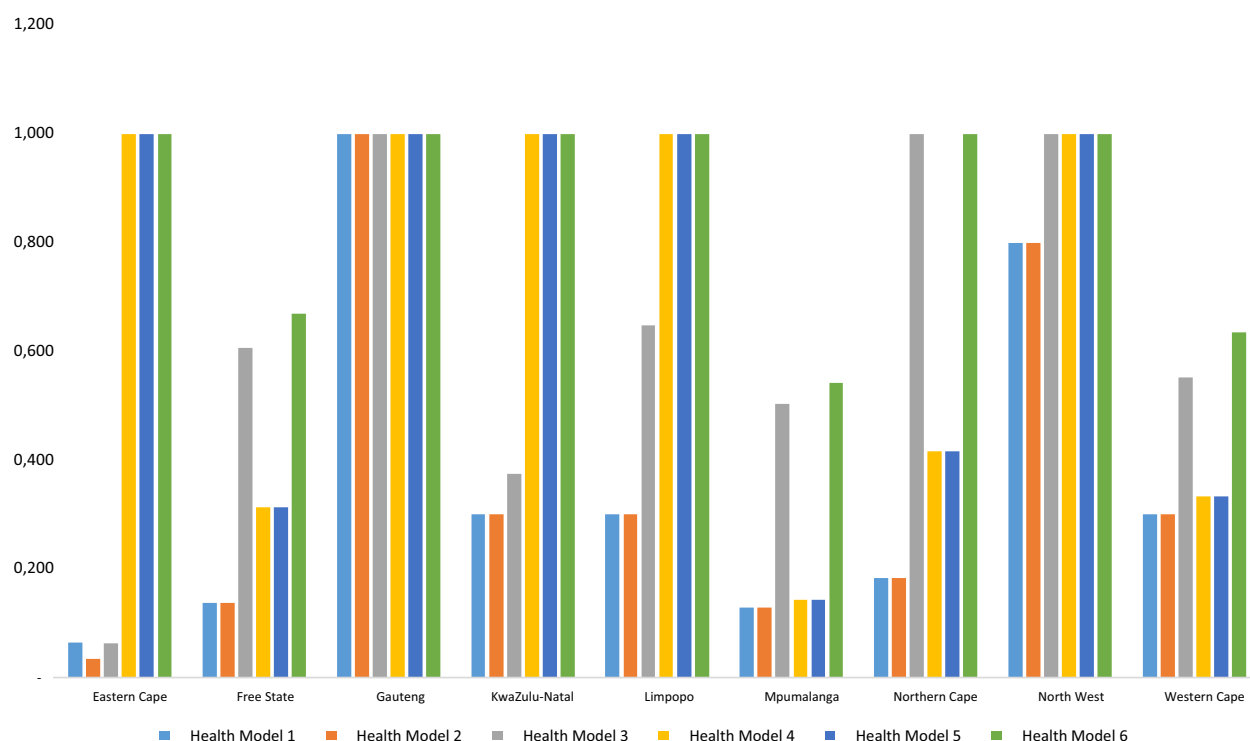


Fig. 7 Summary of provincial health technical efficiency model results. Author's graph based on DEAP 2.1 efficiency results

Received: 27 September 2019 Accepted: 8 January 2020
Published online: 28 January 2020

References

- Afonso A, St. Aubyn M. Relative efficiency of health provision: a DEA approach with non-discretionary inputs. 2006. <https://www.repository.utl.pt/bitstream/10400.5/2655/1/wp332006.pdf>. Accessed 07 Nov 2018.
- Akazili J, Adjui M, Jehu-Appiah C, Zere E. Using data envelopment analysis to measure the extent of technical efficiency of public health centres in Ghana. *BMC Int Health Hum Rights*. 2008;8(1):11.
- Alhassan RK, Nketiah-Amponsah E, Akazili J, Spieker N, Arhinful DK, De Wit TFR. Efficiency of private and public primary health facilities accredited by the national health insurance authority in Ghana. *Cost Eff Resour Alloc*. 2015;13(1):23.
- Anton SG. Technical efficiency in the use of health care resources: a cross-country analysis. *Ann Alexandru Ioan Cuza Univ Econ*. 2013;60(1):1–12.
- Aristovnik A. The impact of ICT on educational performance and its efficiency in selected EU and OECD countries: a non-parametric analysis. 2012. <https://files.eric.ed.gov/fulltext/EJ989207.pdf>. Accessed 14 Nov 2018.
- Asanduloi L, Roman M, Fatulescu P. The efficiency of healthcare systems in Europe: a data envelopment analysis approach. *Procedia Econ Finance*. 2014;10:261–8.
- Asanduloi L, Fatulescu P. The efficiency of EU health using data envelopment analysis. 2014b. <http://relik.vse.cz/2012/sbornik/download/pdf/9-Fatulescu-Puiu-paper.pdf>. Accessed 20 Dec 2018.
- Avkiran NK. Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Econ Plan Sci*. 2001;35(1):57–80.
- Baciu L, Botezat A. A comparative analysis of the public spending efficiency of the new EU member states: a DEA approach. *Emerg Mark Finance Trade*. 2014;50(sup4):31–46.
- Benneyan J, Ceyhan ME, Sunnetci A. Data envelopment analysis of national healthcare systems and their relative efficiencies. In: Elwany MH, Eltawil AB, editors. *Proceedings of the 37th international conference on computers and industrial engineering*; 2007. p. 251–61. <https://pdfs.semanticscholar.org/3814/0ea5c9dd07fb2ef83980c1101169509f5c98.pdf>. Accessed 15 Jan 2019.
- Campanella P, Azzolini E, Izzi A, De Pelone F, La Meo C, Milia CD, Specchia ML, Ricciardi W. Hospital efficiency: how to spend less maintaining quality. *Ann Ist Super Sanita*. 2017;53(1):46–53.
- Chowdhury H, Zelenyuk V, Wodchis W, Laporte A. Efficiency and technological change in health care services in Ontario (No. WP082010). School of Economics, University of Queensland, Australia. 2010. https://www.researchgate.net/profile/Walter_Wodchis/publication/235306670_Efficiency_and_technological_change_in_health_care_services_in_Ontario_An_application_of_Malmquist_Productivity_Index_with_bootstrapping/links/556b4f5f08aefcb861d60ec2/Efficiency-and-technological-change-in-health-care-services-in-Ontario-An-application-of-Malmquist-Productivity-Index-with-bootstrapping.pdf. Accessed 29 Jan 2019.
- Coelli TJ, Rao DSP, O'Donnell CJ, Battese GE. An introduction to efficiency and productivity analysis. New York: Springer Science & Business Media; 2005.
- Competition Commission. Health market enquiry: final findings and recommendations report. Competition Commission. 2019. <https://www.hfassociation.co.za/images/docs/Health-Market-Inquiry-Report.pdf>. Accessed 29 Dec 2019.
- Coovadia H, Jewkes R, Barron P, Sanders D, McIntyre D. The health and health system of South Africa: historical roots of current public health challenges. *Lancet*. 2009;374(9692):817–34.
- Cooper WW, Seiford LM, Tone K. Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software. New York: Springer; 2007.
- Eastern Cape Department of Health. Annual report 2017/18. Eastern Cape Department of Health. 2018. <http://www.ehealth.gov.za/index>

- .php/document-library/annual-reports/send/5-annual-reports/778-dept-of-health-annual-report-2017-18. Accessed 20 Feb 2019.
18. Free State Department of Health. Annual report. 2017/18 Free State Department of Health. Free State Province, South Africa; 2018.
19. Fried HO, Lovell CAK, Schmidt SS. The measurement of productive efficiency and productivity growth. New York: Oxford University Press; 2008.
20. Gannon B. Testing for variation in technical efficiency of hospitals in Ireland. *Econ Soc Rev*. 2005;36(3):273–94.
21. Gauteng Department of Health. Annual report 2017/18. Gauteng Department of Health. 2018. <http://www.gauteng.gov.za/media/Annual%20Reports/2017-2018/Gauteng%20Department%20of%20Health%20Annual%20Report%202017-2018.pdf>. Accessed 20 Feb 2019.
22. Gavurova B, Kocisova K, Belas L, Krajcik V. Relative efficiency of government expenditure on secondary education. *J Int Stud*. 2017;10(2):329–43.
23. Government of the Republic of South Africa. Presidential health summit 2018 report. Government of Republic of South Africa. 2018. https://www.gov.za/sites/default/files/gcis_document/201902/presidential-health-summit-report.pdf. Accessed 06 Apr 2019.
24. Health Systems Trust. South African health review. Health Systems Trust. 2018. <http://www.hst.org.za/publications/South%20African%20Health%20Reviews/SAHR%202018.pdf>. Accessed 28 Apr 2019.
25. Health Systems Trust. South African health review. Health Systems Trust. 2016. <http://www.hst.org.za/publications/South%20African%20Health%20Reviews/SAHR%202016.pdf>. Accessed 28 Apr 2019.
26. Jarjue G, Nor NM, Ghani JA, Jalil SH. Technical efficiency of secondary health care service delivery in the Gambia. *Int J Econ Manag*. 2015;9(1). <http://www.ijem.upm.edu.my/vol9no1/2.%20Technical%20Efficiency%20of%20Secondary%20Health%20Care%20Service%20Delivery%20in%20the%20Gambia.pdf>. Accessed 15 Jan 2019.
27. Kim Y, Kang M. The measurement of health care system efficiency: cross-country comparison by geographical region. 2014. http://s-space.snu.ac.kr/bitstream/10371/91911/1/02_Younhee_Kim.pdf. Accessed 01 Jan 2019.
28. Kirigia JM, Sambo LG, Scheel H. Technical efficiency of public clinics in KwaZulu-Natal Province of South Africa. *East Afr Med J*. 2001;78(3):1–14.
29. Kirigia JM, Lambo E, Sambo LG. Are public hospitals in KwaZulu-Natal Province of South Africa technically efficient? *Afr J Health Sci*. 2000;7(3–4):25.
30. KwaZulu-Natal Department of Health. Annual report 2017/18. KwaZulu-Natal Department of Health. 2018. <http://www.kznhealth.gov.za/2017-2018-Annual-Report.pdf>. Accessed 20 Feb 2019.
31. Lavado R, Domingo G. Public service spending: efficiency and distributional impact—lessons from Asia. 2015. <https://www.econstor.eu/bitstream/10419/128550/1/ewp-435.pdf>. Accessed 20 Jan 2019.
32. Lauro A, Dos Santos Figueiredo OH, Wanke PF. Efficiency of municipal schools in Rio De Janeiro: evidence from two-stage DEA. *J Econ Econ Educ Res*. 2016;17(3):147.
33. Limpopo Department of Health. Annual report: 2017/18. Limpopo Department of Health. 2018. <http://www.doh.limpopo.gov.za/storage/app/media%2FAPP%2FAnnual%20Report%202017-18.pdf>. Accessed 20 Feb 2019.
34. Lo Storto C, Goncharuk AG. Efficiency vs effectiveness: a benchmarking study on European healthcare systems. *Econ Sociol*. 2017;10(3):102–15.
35. Marschall P, Flessa S. Assessing the efficiency of rural health centres in Burkina Faso: an application of data envelopment analysis. *J Public Health*. 2009;17(2):87.
36. Marten R, McIntyre D, Travassos C, Shishkin S, Longde W, Reddy S, Vega J. An assessment of progress towards universal health coverage in Brazil, Russia, India, China, and South Africa (BRICS). *Lancet*. 2014;384(9960):2164–71.
37. Martić M, Novaković M, Baggia A. Data envelopment analysis-basic models and their utilization. *Organizacija*. 2009;42(2):37–43.
38. Masiye F. Investigating health system performance: an application of data envelopment analysis to Zambian Hospitals. *BMC Health Serv Res*. 2007;7(1):58.
39. Mayosi BM, Benatar SR. Health and health care in South Africa—20 years after Mandela. *N Engl J Med*. 2014;371(14):1344–53.
40. McWilliams A, Siegel D, Van Fleet DD. Scholarly journals as producers of knowledge: theory and empirical evidence based on data envelopment analysis. *Organ Res Methods*. 2005;8(2):185–201.
41. Mpumalanga Department of Health. 2017/18 annual report. Mpumalanga Department of Health. Mpumalanga Province, South Africa; 2018.
42. National Treasury. Eastern Cape Province estimates of provincial revenue and expenditure 2018/19. National Treasury. 2018a. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/EC/EC%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
43. National Treasury. Free state provincial treasury, estimates of provincial revenue and expenditure 2018/19. National Treasury. 2018b. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/FS/FS%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
44. National Treasury. Gauteng Province estimates of provincial revenue and expenditure 2018. National Treasury. 2018c. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/GT/GT%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
45. National Treasury. KwaZulu-Natal estimates of provincial revenue and expenditure 2018. National Treasury. 2018d. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/KZN/KZN%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
46. National Treasury. Limpopo provincial government: provincial treasury estimates provincial revenue and expenditure 2018. National Treasury. 2018e. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/LIM/LIM%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
47. National Treasury. Medium term budget policy statement 2018. National Treasury. 2018f. <http://www.treasury.gov.za/documents/mtbps/2018/FullMTBPS.pdf>. Accessed 19 Feb 2019.
48. National Treasury. Mpumalanga provincial government: provincial treasury estimates provincial revenue and expenditure 2018. National Treasury. 2018g. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/MPU/MPU%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
49. National Treasury. Northern Cape provincial government: provincial treasury estimates provincial revenue and expenditure 2018. National Treasury. 2018h. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/NC/NC%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
50. National Treasury. North West provincial government: department of finance estimates provincial revenue and expenditure 2018. National Treasury. 2018i. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/NW/NW%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
51. National Treasury. Western Cape provincial government: provincial treasury estimates provincial revenue and expenditure 2018. National Treasury. 2018j. <http://www.treasury.gov.za/documents/provincial%20budget/2018/3.%20Estimates%20of%20Prov%20Rev%20and%20Exp/WC/WC%20-%20EPRE%20-%20Full%20Document.pdf>. Accessed 19 Feb 2019.
52. Northern Cape Department of Health. Annual report 2017/18. Northern Cape Department of Health. Northern Cape Province, South Africa; 2018.
53. North West Department of Health. Annual report 2017/18. North West Department of Health. North West Province, South Africa; 2018.
54. Prasetyo AD, Zuhdi U. The government expenditure efficiency towards the human development. *Procedia Econ Finance*. 2013;5:615–22.
55. Ramírez Hassan A. Consequences of omitting relevant inputs on the quality of the data envelopment analysis under different input correlation structures. 2008. <http://repository.eafit.edu.co/xmlui/bitstream/handle/10784/556/2008-1%20Andres%20Ramirez%20Hassan.pdf?sequence=1&isAllowed=y>. Accessed 15 Nov 2018.
56. Statistics South Africa. Gross domestic product, quarter 4 2017. Statistics South Africa. 2017. http://www.statssa.gov.za/publications/P0441/GDP_Q4_2017_Media_presentation.pdf. Accessed 23 Mar 2019.
57. Statistics South Africa. Community survey 2016 statistical release. Statistics South Africa. 2016. <http://cs2016.statssa.gov.za/wp-content/>

- uploads/2016/07/NT-30-06-2016-RELEASE-for-CS-2016-_Statistical-relea_s_1-July-2016.pdf. Accessed 20 July 2016.
58. Taylor B, Harris G. Relative efficiency among South African universities: a data envelopment analysis. *High Educ*. 2004;47(1):73–89.
 59. Varabyova Y, Schreyögg J. International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches. *Health Policy*. 2013;112(1–2):70–9.
 60. Verhoeven M, Gunnarsson V, Carcillo S. Education and health in G7 countries: achieving better outcomes with less spending (No. 2007-2263). International Monetary Fund. 2007. https://www.researchgate.net/profile/Stephane_Carcillo/publication/263084613_WP07263_Education_and_Health_in_G7_Countries_Achieving_Better_Outcomes_with_Less_Spending/links/00b49539c08ba4a556000000/WP-07-263-Education-and-Health-in-G7-Countries-Achieving-Better-Outcomes-with-Less-Spending.pdf. Accessed 01 Apr 2019.
 61. Wang EC, Alvi E. Relative efficiency of government spending and its determinants: evidence from East Asian countries. *Eurasian Econ Rev*. 2011;1(1):3–28.
 62. Western Cape Government Health. Health 2017/18 annual report. Western Cape Government Health. 2018. https://www.westerncape.gov.za/sites/www.westerncape.gov.za/files/wcgh_annual_report_2018.pdf. Accessed 20 Feb 2019.
 63. Yawe BL. Technical efficiency and productivity of primary schools in Uganda. 2014. <https://www.africaportal.org/documents/17554/RP277.pdf>. Accessed 31 Jan 2019.
 64. Yuan Y, Shan M. The educational efficiency evaluation framework: by using DEA model and CA method. *Int J Inf Educ Technol*. 2016;6(12):923–6.
 65. Zere E, Mbeeli T, Shangula K, Mandlhate C, Mutirua K, Tjivambi B, Kapenambili W. Technical efficiency of district hospitals: evidence from Namibia using data envelopment analysis. *Cost Eff Resour Alloc*. 2006;4(1):1.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

