Mongolian Journal of Economic Review-MJER **Volume 30/19 (2026)**
ARTICLE
https://doi.org/10.69588/mjer.v30i19.5654

Data Envelopment Analysis Beyond Hospitals: Efficiency Measurement in Primary Care, District Health Systems, and Long-Term Care

Ganzorig Batbaatar [1,2§], Buyantogtokh Batsukh [2§], Amarzaya Batchimed [2], Byambakhorloo Sukhbaatar [2], Byambatseren Boldbaatar [3], Nursabi Khizatkhan [3] Dariimaa Ganbat [3*]

[1] *National Center for Pathology of Mongolia, Ulaanbaatar, Mongolia;*
[2] *Graduate School, University of Finance and Economics, Ulaanbaatar, Mongolia;*
[3] *Graduate School, Mongolian National University of Medical Sciences, Ulaanbaatar, Mongolia*
[§] *Equally contributed*

Corresponding author:* **Dariimaa Ganbat E-mail: **dariimaa@mnums.edu.mn**

Received: 28 February 2026 | Accepted: 21 May 2026 | Published online: 29 June 2026

Kye words: *Technical efficiency; primary care; long-term care; district health system; decision-making unit; resource allocation; public-sector management*

Abstract: *Data envelopment analysis (DEA) is widely used to assess hospital efficiency, but its application in primary care, district health systems, and long-term care remains methodologically inconsistent.* ***Objective:*** *This review synthesized non-hospital DEA applications, focusing on decision-making units, input–output selection, quality measurement, model specification, and implications for public-sector resource allocation.* ***Methods:*** *Following PRISMA 2020 principles, several electronic databases were used. Eligible studies applied DEA to primary care, district/regional health systems, or long-term care and reported decision-making units, inputs, outputs, and model specifications.* ***Results:*** *Thirty studies generated 33 setting-specific observations. Mean technical efficiency was 0.593 in primary care, 0.608 in long-term care, and 0.458 in district/regional systems.* ***Conclusion:*** *Non-hospital health systems show substantial efficiency gains, but policy interpretation should be cautious due to limited comparability, volume-based outputs, and weak quality adjustment. This review offers methodological guidance for context-sensitive DEA benchmarking in resource-constrained public health systems.*

https://doi.org/10.69588/mjer.v30i19.5654

I. INTRODUCTION

Health systems across the world are under growing pressure to use limited resources more wisely while maintaining equitable access, acceptable quality, and responsiveness to population needs. This challenge is particularly important in resource-constrained settings, where even small inefficiencies can affect service coverage and health outcomes (World Health Organization, 2010). Data envelopment analysis (DEA) has become one of the most widely used approaches for assessing relative efficiency in health care because it can evaluate multiple inputs and outputs simultaneously. However, much of the existing DEA literature has focused on hospitals, while primary care, district health systems, community-based services, and long-term care have received comparatively less attention (Hollingsworth, 2008; Kohl et al., 2019).

This review focuses on the methodological and policy issues that arise when DEA is applied to non-hospital health care settings. In particular, it examines how decision-making units (DMUs) are defined, how service outputs and quality are measured, and how model choices influence the interpretation of efficiency results. These issues are important because poorly defined DMUs, narrow output measures, or inappropriate model assumptions may lead to misleading conclusions and weak policy recommendations. A clearer understanding of these methodological challenges is therefore essential for using DEA results to guide resource allocation and performance improvement in public-sector health systems (Dyson et al., 2001; Jacobs, Smith, & Street, 2006). This systematic review addresses three main research questions. First, how are DMUs defined in non-hospital health care settings, and what comparability problems arise because of differences in organizational structure, service scope, and regulatory context? Second, what types of output measures are used, and to what extent is quality of care incorporated into efficiency assessment? Third, which DEA model specifications, including constant or variable returns to scale, input or output orientation, quality adjustment, two-stage regression, and bootstrap inference, are most suitable for these settings, and how should efficiency scores be interpreted for management and policy purposes?

By synthesizing evidence setting-specific applications, this review provides practical methodological guidance for researchers, health economists, public managers, and policymakers who aim to use DEA for efficiency assessment in primary care, district health systems, and long-term care.

Theoretical and Methodological Foundations of DEA

DEA is a non-parametric linear programming method that constructs an empirical efficiency frontier from observed combinations of inputs and outputs. Each DMU is then compared with this frontier to estimate its relative efficiency (Charnes, Cooper, & Rhodes, 1978; Cooper, Seiford, & Tone, 2007). In simple terms, a unit is considered technically efficient when no other observed unit, or combination of units, can produce more outputs with the same or fewer inputs, or produce the same outputs using fewer resources (Farrell, 1957). One of the main advantages of DEA is that it does not require price information or a pre-specified production function. This makes it especially useful in public and nonprofit health care settings, where market prices are often absent and services are delivered with multiple objectives rather than profit maximization (Worthington, 2004). The two classical DEA models are the CCR model, which assumes constant returns to scale, and the BCC model, which allows variable returns to scale and separates pure technical efficiency from scale efficiency (Charnes, Cooper, & Rhodes, 1978; Banker, Charnes, & Cooper, 1984).

https://doi.org/10.69588/mjer.v30i19.5654

In non-hospital settings, the assumption of constant returns to scale is often difficult to justify. Primary care facilities, community health centers, district health offices, and long-term care institutions may differ considerably in size, function, population coverage, and available resources. For this reason, BCC or variable returns-to-scale models are often more appropriate, particularly when the units being compared operate under different scale conditions (Giuffrida & Gravelle, 2001). DEA models may also be input-oriented or output-oriented. In public health systems, where budgets and staffing levels are often fixed in the short term, an output-oriented approach may be preferable because it asks how much service delivery could be increased using existing resources (Moran, Suhrcke, & Nolte, 2023). Several extensions of standard DEA have been developed to address more complex analytical questions. Slack-based measure models can identify specific input excesses or output shortfalls more directly (Tone, 2001). Malmquist productivity analysis can assess changes in productivity over time by decomposing change into technical efficiency change, technological change, and scale-related change (Färe, Grosskopf, Norris, & Zhang, 1994). Two-stage DEA methods can also be used to examine whether contextual or environmental factors explain variation in efficiency after the initial efficiency scores have been estimated (Simar & Wilson, 2007).

Conceptual Challenges in Non-Hospital Efficiency Measurement

Applying DEA outside hospital settings creates several methodological difficulties. The first challenge is DMU comparability. Hospitals usually have more standardized organizational structures, but non-hospital units are often much more diverse. For example, primary care clinics, district health systems, nursing homes, and community health programs may differ in governance, staffing, catchment population, service responsibilities, financing arrangements, and local infrastructure. As a result, measured inefficiency may partly reflect structural differences rather than true differences in managerial or technical performance (Dyson et al., 2001; Kringos et al., 2010). The second challenge concerns output measurement. Non-hospital health care often involves prevention, health promotion, chronic disease management, rehabilitation, community outreach, functional support, and continuity of care. These activities are not fully captured by simple volume-based outputs such as visits, consultations, admissions, or bed-days (World Health Organization, 2008; Donabedian, 1988; Kane et al., 2007). If DEA models rely only on service volume, they may unintentionally reward units that produce more activities but deliver lower-quality care. For this reason, quality adjustment is especially important in non-hospital settings. Relevant indicators may include patient-reported outcomes, clinical process measures, avoidable complications, continuity of care, safety indicators, or composite quality measures (Hussey et al., 2009).

The third challenge is contextual heterogeneity. Efficiency scores may be influenced by factors outside the control of service managers, including geography, population density, socioeconomic conditions, financing rules, regulatory arrangements, infrastructure, disease burden, case-mix, and health information capacity. If these factors are ignored, DEA may unfairly classify disadvantaged units as inefficient or overestimate the performance of units operating in more favorable environments (Simar & Wilson, 2007; Kruk et al., 2018).

Policy Relevance and Interpretation of DEA Findings

From a health economics and public management perspective, DEA can support policy in three main ways. First, it can be used for benchmarking by identifying relatively efficient units and estimating the potential improvement required among lower-performing units

https://doi.org/10.69588/mjer.v30i19.5654

(Hollingsworth & Peacock, 2008). Second, it can inform resource allocation by linking efficiency evidence with decisions about budgets, staffing, infrastructure, and service expansion (Smith & Street, 2005). Third, DEA can support management reform by identifying operational weaknesses, guiding process improvement, and strengthening accountability systems (Bevan & Hood, 2006). However, DEA results should not be interpreted mechanically. Efficiency is a relative measure, meaning that a unit is efficient only in relation to the observed sample and the constructed frontier, not according to an absolute standard of best practice (Bogetoft & Otto, 2011). A low efficiency score may indicate managerial underperformance, but it may also reflect unmeasured differences in population need, case-mix complexity, service quality, geographic barriers, or institutional constraints (Simar & Wilson, 2007). Therefore, DEA-based recommendations should be used carefully and combined with contextual information, qualitative evidence, and equity considerations. In public-sector health systems, the goal should not simply be to maximize output volume, but to improve the fair, efficient, and high-quality use of available resources (Cylus, Papanicolas, & Smith, 2016).

II. METHODOLOGY

2.1 Review Design, Protocol, and Eligibility Criteria

This systematic review was conducted and reported in accordance with the PRISMA 2020 guidelines to ensure transparency and completeness in the review process (Page et al., 2021). A protocol was developed to guide the search strategy, study selection, and data extraction procedures; however, it was not prospectively registered in an international review registry. This is acknowledged as a methodological limitation, as protocol registration helps reduce the risk of selective reporting and improves review transparency.

Eligible studies were selected using predefined inclusion and exclusion criteria. These criteria were designed to identify empirical studies that applied Data Envelopment Analysis (DEA) to assess efficiency in non-hospital health care settings, including primary care, district or regional health systems, community-based services, nursing homes, and long-term care facilities. The detailed eligibility criteria are summarized in **Table 1**.

Table 1. Eligibility criteria for study inclusion

Criterion	Inclusion criteria	Exclusion criteria
Population/ setting	Primary care facilities, community health centers, district or regional health systems, long-term care facilities, nursing homes, aged-care institutions, residential care facilities, or home-based elderly-care services.	Hospital-only studies, national macro-health-system studies without a primary care, district/regional, or long-term care focus, and studies outside the healthcare sector.
Study design/ method	Empirical studies applying DEA or DEA extensions, including CCR, BCC, SBM, bootstrap DEA, super-efficiency, two-stage DEA, three-stage DEA, order-m, or Malmquist productivity analysis.	Studies using only stochastic frontier analysis, cost-effectiveness analysis, regression-only productivity models, or conceptual discussion without DEA estimates.
Reporting detail	Studies reporting the decision-making unit definition, input variables, output variables, DEA model specification, and interpretable efficiency results.	Studies lacking sufficient information on DMU definition, input/output variables, DEA specification, or efficiency results.
Publication type and language	Peer-reviewed journal articles and full research papers published in English.	Conference abstracts, editorials, letters, commentaries, dissertations without peer review, non-peer-reviewed reports, and duplicate publications.

https://doi.org/10.69588/mjer.v30i19.5654

Publication period	Studies published from database inception to 31 January 2026.	Studies published after the final search update.
Outcome of interest	DEA efficiency score, technical efficiency, scale efficiency, productivity change, or setting-specific methodological insight relevant to DEA interpretation.	Studies reporting no efficiency score and no methodological information relevant to DEA.

2.2 Information Sources, Study Selection, and Data Extraction

A comprehensive literature search was conducted to identify studies applying Data Envelopment Analysis (DEA) to efficiency assessment in non-hospital health care settings. Six electronic sources were searched: Web of Science, Scopus, EMBASE, EconLit, Google Scholar, and PubMed. The search strategy combined Boolean operators and broader semantic terms to capture studies related to DEA, efficiency measurement, and frontier-based analysis in primary care, community health, district health systems, regional health services, nursing homes, long-term care, elderly care, and residential care settings. The main search terms included methodological keywords such as “data envelopment analysis,” “DEA,” “efficiency measurement,” “technical efficiency,” and “frontier analysis,” combined with setting-related terms including “primary care,” “primary health care,” “health center,” “community health,” “district health,” “regional health system,” “nursing home,” “long-term care,” “elderly care,” and “residential care.” The full search strategy is provided in the supplementary material. The initial search was conducted in December 2025 and was updated on 31 January 2026. In addition, the reference lists of included studies and relevant review articles were manually screened to identify further eligible publications.

Study selection was carried out independently by two reviewers using predefined eligibility criteria. Titles and abstracts were first screened to exclude clearly irrelevant records. Full-text articles were then retrieved and assessed for studies that appeared to meet the inclusion criteria or required further evaluation. Any disagreements between reviewers were resolved through discussion, and a third reviewer was consulted when consensus could not be reached. The overall selection process was documented using a PRISMA flow diagram. Data extraction was performed using a standardized extraction form developed for this review. The extracted information included general study characteristics, such as author, publication year, country, and type of health care setting; the definition and number of decision-making units (DMUs); and relevant organizational or geographic characteristics. Details on DEA model specifications were also recorded, including input and output variables, returns-to-scale assumptions, model orientation, quality adjustment, two-stage regression, and the use of bootstrap or other statistical inference methods.

Input variables commonly included labor resources, such as physicians, nurses, and other health personnel; capital resources, such as beds, equipment, and facilities; and operating expenditures. Output variables included service volume indicators, such as consultations, visits, admissions, and bed-days; intermediate outputs, such as vaccinations and screening activities; and, where available, health outcomes or quality indicators, including mortality, morbidity, quality of life, and patient care quality measures. Reported efficiency scores, including mean, median, range, and distribution, were also extracted. Data extraction was completed independently by two reviewers and cross-checked to ensure accuracy and consistency.

https://doi.org/10.69588/mjer.v30i19.5654

2.3 Quality appraisal

The assessment of reporting quality utilized a modified framework that amalgamates the Joanna Briggs Institute (JBI) critical appraisal checklist for economic evaluations and Drummond's economic analysis checklist. This framework encompassed criteria such as the clarity of the research question, transparency in defining the decision-making unit (DMU), thorough measurement of inputs and outputs, and appropriate handling of contextual variables, among others. Each criterion received a score of "yes," "partial," or "no," with the quality assessment serving to enhance the interpretation of findings without excluding studies based on their quality scores.

2.4 Data synthesis

Due to substantial heterogeneity in DMU definitions, input–output specifications, and contextual factors, conventional meta-analysis of efficiency scores was not appropriate (Higgins et al., 2003). Instead, we conducted descriptive synthesis:

1. Narrative synthesis: Summarized DMU definitions, input–output choices, model specifications, and quality integration approaches by setting type.
2. Quantitative summary: Calculated pooled mean efficiency scores and 95% bootstrap confidence intervals for each setting type (primary care, district health systems, long-term care) using 10,000 bootstrap resamples (Simar & Wilson, 1998).
3. Comparative analysis: Compared methodological approaches and efficiency patterns across settings, highlighting implications for policy interpretation.

Bootstrap confidence intervals were used to quantify uncertainty around pooled means without assuming normality or independence of efficiency scores (Simar & Wilson, 1998).

III. RESULTS

3.1 Study selection

The search identified 920 records. After removing duplicates, 670 unique records were screened by title and abstract, yielding 250 potentially eligible studies. Full-text review resulted in 30 studies meeting all inclusion criteria. Several studies reported efficiency estimates for multiple settings (e.g., separate analyses for primary care and district health systems), generating 33 setting-specific observations: 23 for primary care, 7 for long-term care, and 3 for district/regional health systems. The PRISMA flow diagram is presented in **Figure 1**.

3.2 Study characteristics

Thirty studies met the inclusion criteria and generated 33 setting-specific DEA observations because some studies evaluated more than one healthcare setting. Each setting-specific DEA application was therefore treated as a separate analytical unit. The included studies were published between 2002 and 2026, with most published after 2015 (n=22, 73%). They represented high-income countries (n=18, 60%), upper-middle-income countries (n=8, 27%), and low- or lower-middle-income countries (n=4, 13%). Primary care accounted for most observations (n=23), followed by long-term care (n=7) and district/regional health systems (n=3).

https://doi.org/10.69588/mjer.v30i19.5654

3.3 Decision-making unit definition and comparability

Primary care: DMUs included individual health centers (n=15), community health posts (n=4), primary care practices (n=3), and municipal primary care networks (n=1). Sample sizes ranged from 12 to 2,842 DMUs (median: 87). Organizational heterogeneity was substantial: some studies compared solo physician practices with multi-disciplinary health centers; others pooled urban and rural facilities despite differences in service scope, staffing, and patient populations (Giuffrida, 1999) (Zavras, Tsakos, Economou, & Kyriopoulos, 2002) (Su, Du, Fan, & Wang, 2023).

Long-term care: DMUs were predominantly nursing homes (n=5) or residential care facilities (n=2). Sample sizes ranged from 30 to 555 DMUs (median: 120). Heterogeneity arose from variation in ownership (public, private non-profit, private for-profit), size (small homes vs. large facilities), and case-mix (independent living vs. high-dependency care) (Shimshak, Lenard, & Klimberg, 2009) (Shao, Yu, & Feng, 2021).

District/regional systems: DMUs were administrative or geographic aggregates (n=3), such as district health offices coordinating multiple facilities and public health programs. Sample sizes were small (range: 12–30 DMUs). These studies faced the greatest comparability challenges, as districts varied widely in population size, geographic area, infrastructure, and service integration (Rahimi et al., 2021) (Yitbarek et al., 2021).

https://doi.org/10.69588/mjer.v30i19.5654

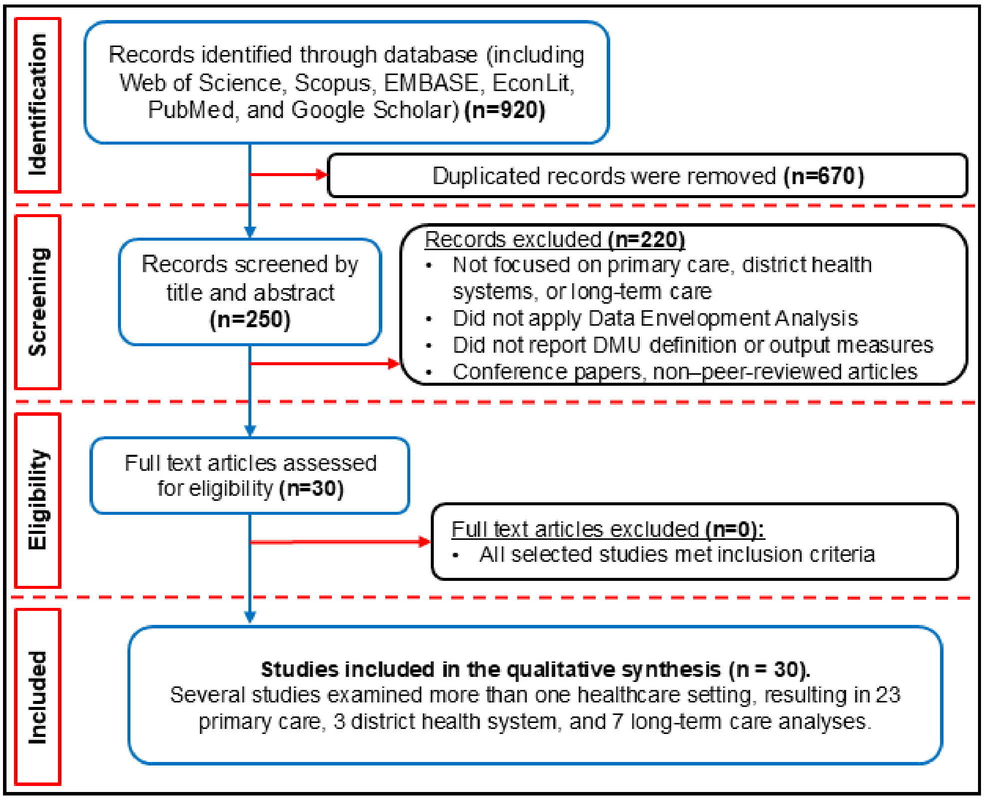

Figure 1. PRISMA flow diagram for study identification, screening, eligibility, and inclusion.

3.4 Input and output measurement

Inputs: Labor inputs were reported in all studies, typically as full-time equivalent (FTE) counts of physicians, nurses, and support staff. Capital inputs were reported in 18 studies (60%), including number of beds, examination rooms, or equipment. Operating expenditure was used as a composite input in 9 studies (30%). Few studies adjusted inputs for skill mix, experience, or training level (Giuffrida & Gravelle, 2001). Outputs: Service volume measures dominated: consultations or visits (n=20, 67%), admissions or discharges (n=8, 27%), bed-days (n=6, 20%), and preventive services such as vaccinations or screenings (n=7, 23%). Intermediate health outcomes (e.g., disease detection rates, immunization coverage) were reported in 5 studies (17%). Final health outcomes (e.g., mortality, morbidity, quality of life) were reported in only 3 studies (10%) (Moran, Suhrcke, & Nolte, 2023) (Moran, Suhrcke, & Nolte, 2023).

Quality measurement: Quality indicators were incorporated in 12 studies (40%), using approaches including:

- Quality-adjusted outputs: Weighting service volumes by patient satisfaction scores, clinical process indicators, or accreditation status (n=6) (Clement et al., 2008) (McGarvey, Thorsen, Thorsen, & Hozain, 2019).
- Undesirable outputs: Treating adverse events, readmissions, or complaints as undesirable outputs in slacks-based models (n=3) (Tone & Tsutsui, 2010).
- Two-stage quality adjustment: Regressing efficiency scores on quality indicators in a second stage (n=3) (Chilingerian & Sherman, 2004).

https://doi.org/10.69588/mjer.v30i19.5654

However, 18 studies (60%) relied solely on volume-based outputs without explicit quality adjustment, raising concerns about potential trade-offs between throughput and care quality (Nayar & Ozcan, 2008).

3.5 Data envelopment analysis model specifications

Returns to scale: Variable returns to scale (VRS) models were used in 22 studies (73%), constant returns to scale (CRS) in 5 studies (17%), and both CRS and VRS in 3 studies (10%). VRS was more common in primary care and long-term care, reflecting recognition that small facilities may operate below optimal scale (Banker, Conrad, & Strauss, 1986).

Orientation: Output orientation was used in 18 studies (60%), input orientation in 8 studies (27%), and non-oriented models (e.g., slacks-based measures) in 4 studies (13%). Output orientation was preferred in public health systems with fixed budgets and high service demand (Coelli, Rao, O'Donnell, & Battese, 2005).

Advanced methods: Two-stage DEA, regressing efficiency scores on contextual variables (e.g., population density, socioeconomic status, case-mix), was applied in 11 studies (37%) (Simar & Wilson, 2007) (Afonso & St. Aubyn, 2006). Bootstrap DEA, providing confidence intervals for efficiency scores, was used in 4 studies (13%) (Simar & Wilson, 1998) (Simar & Wilson, 2000). Malmquist productivity indices, decomposing efficiency change over time, were reported in 3 studies (10%) (Färe, Grosskopf, Norris, & Zhang, 1994).

3.6 Efficiency scores

The pooled efficiency estimates differed across healthcare settings, with long-term care showing the highest mean score, district/regional systems the lowest mean score, and primary care showing substantial variation across observations (**Table 2**).

Table 2. Pooled mean technical efficiency scores by setting type

Setting Type	Number of Observations	Pooled Mean Efficiency	95% Bootstrap CI	Range
Primary care	23	0.593	0.482–0.705	0.312–0.891
Long-term care	7	0.608	0.426–0.791	0.401–0.823
District/regional systems	3	0.458	0.151–0.765	0.312–0.587

Primary care: Mean efficiency was 0.593 (95% CI: 0.482–0.705), indicating that primary care facilities could, on average, increase outputs by approximately 69% ($1/0.593 \approx 1.69$) with existing inputs if they operated at the efficiency frontier. Efficiency varied widely across studies (range: 0.312–0.891), reflecting differences in country context, DMU definition, and model specification (Ramirez-Valdivia, Maturana-Ross, & Salvo-Garrido, 2015) (Purohit, 2016).

https://doi.org/10.69588/mjer.v30i19.5654

Long-term care: Mean efficiency was 0.608 (95% CI: 0.426–0.791), slightly higher than primary care but with substantial uncertainty due to the smaller number of studies. Efficiency ranged from 0.401 to 0.823 (Shimshak, Lenard, & Klimberg, 2009) (Cheng et al., 2015).

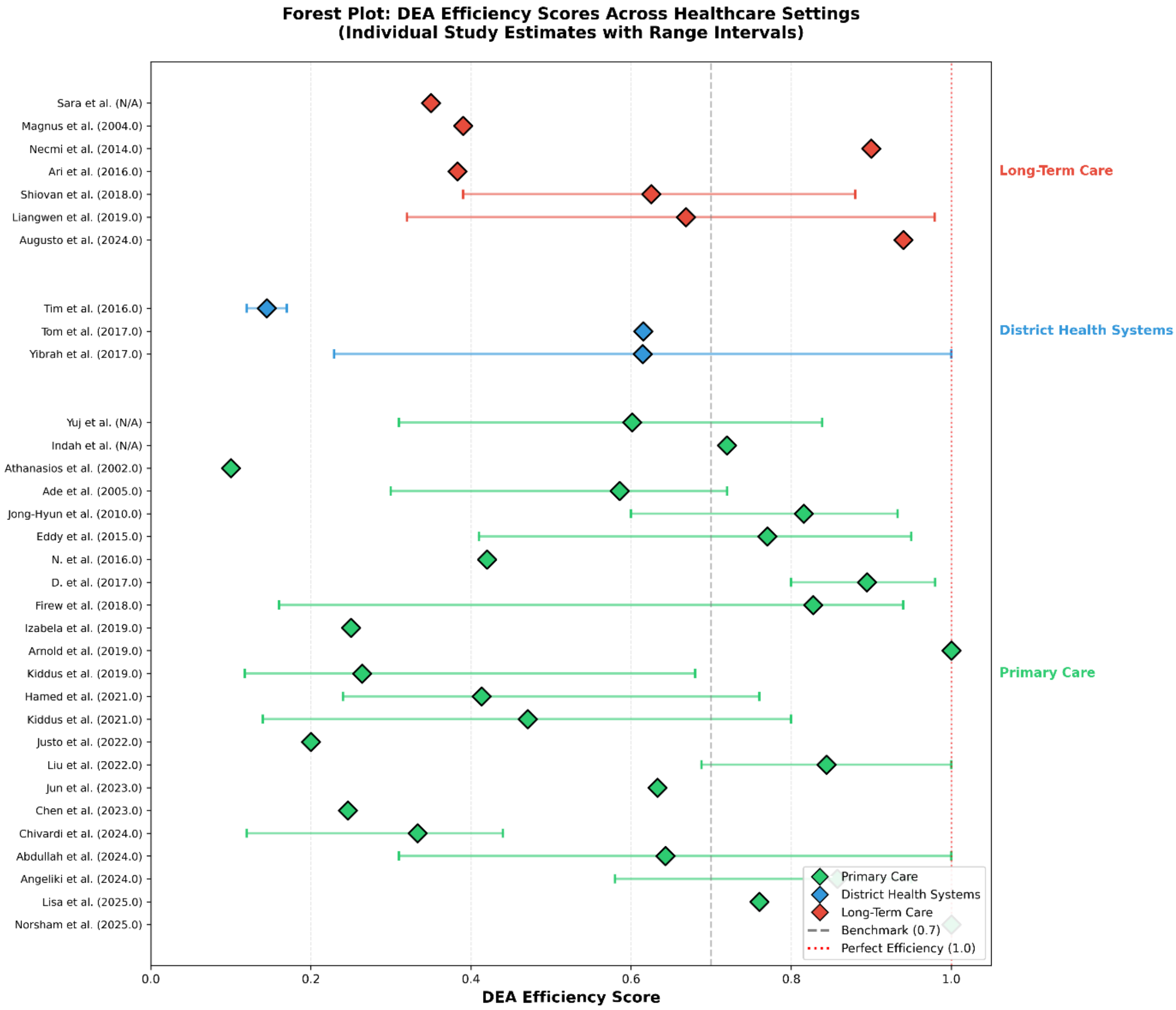

Figure 2. Forest plot of pooled DEA efficiency scores by healthcare setting.

District/regional systems: Mean efficiency was 0.458 (95% CI: 0.151–0.765), the lowest of the three settings, with very wide confidence intervals reflecting both low efficiency and high uncertainty. Only three studies examined district systems, limiting generalizability (Rahimi et al., 2021) (Yitbarek et al., 2021) (Renner & Kirigia, 2005). The pooled mean efficiency estimates and their bootstrap confidence intervals are summarized visually in **Figure 2**. All pooled estimates fell below the conventional 0.70 benchmark often cited in DEA literature (Sherman, 1984), suggesting substantial improvement potential. However, the wide confidence intervals and high heterogeneity caution against direct policy inference without context-sensitive interpretation.

https://doi.org/10.69588/mjer.v30i19.5654

To illustrate the dispersion and within-setting variation of individual DEA efficiency estimates, **Figure 3** presents violin plots with study-level observations across primary care, district health systems, and long-term care. Primary care studies showed a mean efficiency score of 0.593 (95% CI: 0.482-0.705) and substantial dispersion, indicating considerable between-study variation and several estimates below the pragmatic 0.70 benchmark. District/regional health systems had the lowest mean efficiency score of 0.458 (95% CI: 0.151-0.765), but this estimate should be interpreted cautiously because it was based on only three observations and had wide uncertainty. Long-term care showed a mean efficiency score of 0.608 (95% CI: 0.426-0.791), with moderate dispersion and substantial variation across studies.

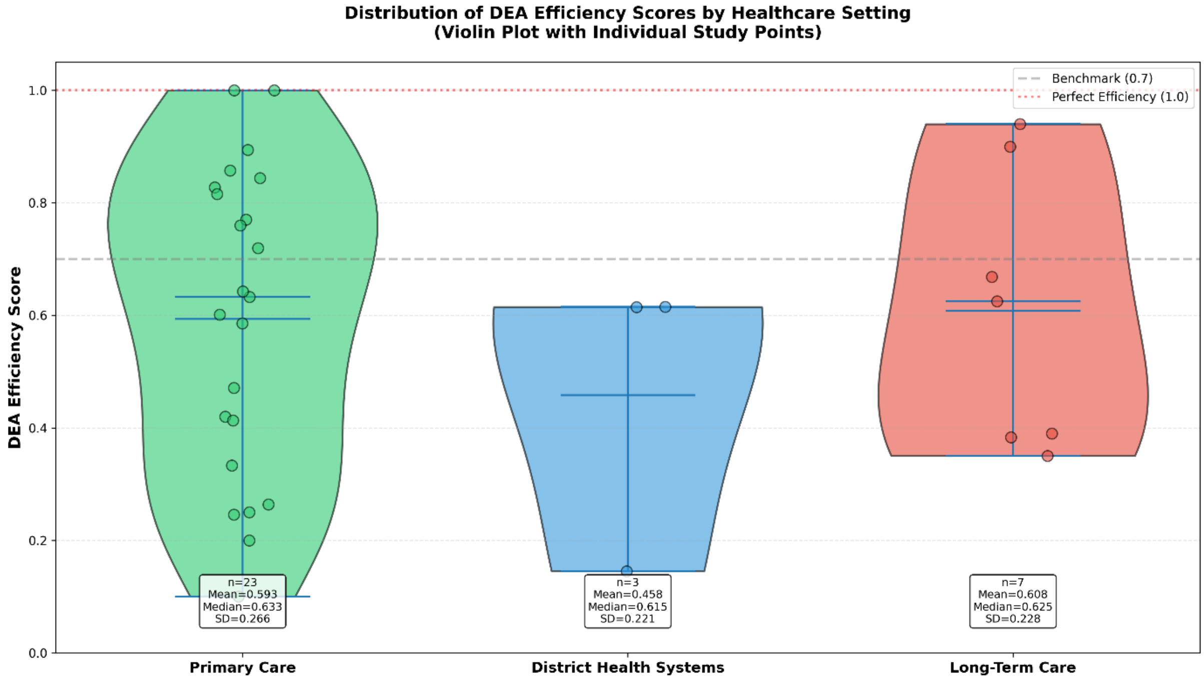

Figure 3. Distribution of DEA efficiency scores by healthcare setting.

3.7 Quality appraisal results

Quality appraisal results showed a clear pattern of strengths and limitations. Most studies (n=26, 87%) clearly stated research questions and policy relevance. DMU definition and comparability were transparently described in 20 studies (67%) but inadequately addressed in 10 studies (33%), particularly those pooling heterogeneous organizational types. Input measurement was comprehensive in 24 studies (80%), but output measurement was less robust: only 12 studies (40%) incorporated quality indicators, and only 3 studies (10%) measured final health outcomes. DEA model specification was explicitly justified in 18 studies (60%). Two-stage regression or environmental adjustment was applied in 11 studies (37%). Sensitivity analysis or bootstrap inference was reported in 7 studies (23%). Interpretation acknowledged limitations in 22 studies (73%).

https://doi.org/10.69588/mjer.v30i19.5654

IV. DISCUSSION

4.1 Principal findings

This systematic review examined DEA applications in primary care, district health systems, and long-term care, revealing three key findings:

Significant efficiency disparities: Average efficiency scores ranged from 0.458 (district systems) to 0.608 (long-term care), all below standard benchmarks, suggesting a possible output increase or input decrease of 64% to 118% if DMUs operated at peak efficiency. Methodological diversity: Variations in DMU definitions, input-output specifications, and model selections hindered the comparability and applicability of efficiency assessments. Insufficient quality integration: Merely 40% of studies included quality metrics, which raises concerns about the potential trade-off between efficiency improvements and care quality.

4.2 Methodological implications

DMU comparability: The legitimacy of efficiency measurement based on Data Envelopment Analysis (DEA) is fundamentally contingent upon the comparability of Decision-Making Units (DMUs) (Dyson et al., 2001). The juxtaposition of a single physician's practice with a multi-disciplinary health facility, or a small rural health post against a large urban clinic, conflates organizational structure and service range with technical efficiency. Subsequent investigations should categorize DMUs by type of organization, extent of services provided, and regulatory context, or implement meta-frontier analysis to distinguish group-specific efficiency from technological disparities (O'Donnell, Rao, & Battese, 2008).

Output measurement and quality: Dependence exclusively on volume-based outputs (such as consultations, visits, or bed-days) without quality adjustments may inadvertently promote high throughput while compromising care quality (Hussey et al., 2009). Quality-adjusted DEA methodologies, which weight outputs according to patient satisfaction, clinical process measures, or health outcomes, are methodologically sound and ought to be regarded as standard practice (Clement et al., 2008) (McGarvey, Thorsen, Thorsen, & Hozain, 2019). Nevertheless, quality data frequently remain inaccessible or untrustworthy in resource-limited environments, underscoring the imperative for enhanced health information systems (AbouZahr & Boerma, 2005). Model specification: The determination of whether to use Constant Returns to Scale (CRS) or Variable Returns to Scale (VRS), as well as the choice between input and output orientation, must be explicitly rationalized in relation to the institutional context. VRS is suitable when DMUs function at differing scales and when scale efficiency is of interest (Banker, Charnes, & Cooper, 1984). Output orientation is pertinent in scenarios where budgets are fixed and service demand surpasses supply (Coelli, Rao, O'Donnell, & Battese, 2005). Two-stage DEA, which regresses efficiency scores on contextual factors, can differentiate managerial inefficiency from environmental limitations but necessitates meticulous attention to statistical inference (Simar & Wilson, 2007).

Uncertainty quantification: Merely 23% of studies have disclosed sensitivity analyses or bootstrap confidence intervals. Efficiency scores represent point estimates that are vulnerable to sampling variability and uncertainties in model specification (Simar & Wilson, 1998). Bootstrap DEA offers confidence intervals that elucidate this uncertainty and should be consistently presented (Simar & Wilson, 2000).

https://doi.org/10.69588/mjer.v30i19.5654

4.3 Policy implications

Context-sensitive benchmarking: DEA efficiency scores constitute relative metrics rather than definitive benchmarks. A Decision-Making Unit (DMU) is deemed efficient only in relation to the observed efficiency frontier, which may itself be less than optimal (Bogetoft & Otto, 2011). Low efficiency may indicate managerial inadequacies; however, it may also be indicative of unmeasured complexities in case-mix, disparities in quality, or environmental limitations such as rural positioning, insufficient infrastructure, or ineffective governance (Simar & Wilson, 2007). Consequently, policy suggestions arising from DEA analyses must be sensitive to context, corroborated with qualitative data, and executed with a focus on equity and accessibility concerns (Cylus, Papanicolas, & Smith, 2016).

Resource allocation and public-sector management: DEA can enhance resource distribution by pinpointing exemplary DMUs and assessing the potential for improvement in those that are underperforming (Hollingsworth & Peacock, 2008). Nevertheless, the conversion of efficiency scores into budgetary allocations necessitates careful consideration. Imposing penalties on low-efficiency DMUs through budget reductions may intensify resource limitations and degrade performance; conversely, rewarding high-efficiency DMUs without comprehending the underlying factors of their efficiency could lead to resource misallocation (Smith & Street, 2005). A more sophisticated strategy correlates efficiency assessment with capacity enhancement, process optimization, and management assistance (Bevan & Hood, 2006).

Relevance for resource-constrained systems: In public health systems constrained by resources, such as Mongolia and analogous contexts, DEA serves as a pragmatic instrument for assessing efficiency when pricing information is absent and multiple inputs and outputs need simultaneous evaluation (Worthington, 2004). However, the methodological prerequisites, including comparable DMUs, quality-adjusted outputs, and context-sensitive interpretation, are rigorous. Enhancing health information infrastructures, investing in data integrity, and fostering analytical capabilities are essential for effective DEA-driven performance monitoring and resource allocation (AbouZahr & Boerma, 2005).

4.4 Limitations and Future Research Directions

This review has several limitations that should be considered when interpreting the findings. First, the review protocol was not registered prospectively, which may increase the risk of selective reporting. Second, substantial heterogeneity was observed across the included studies, particularly in the definitions of decision-making units (DMUs), the selection of input and output variables, and the use of contextual factors. Because of this variability, a conventional meta-analysis was not appropriate. Therefore, the aggregated efficiency scores should be interpreted as descriptive summaries rather than as causal or directly comparable estimates. Third, only a small number of studies focused specifically on district or regional health systems. This limits the generalizability of the findings to decentralized health systems or lower-level service delivery settings. Fourth, publication bias may have influenced the results, as studies reporting low efficiency scores or methodological challenges may be less likely to be published. As a result, the reported efficiency estimates may overstate actual performance. Finally, the review relied mainly on descriptive synthesis rather than meta-regression or formal subgroup analysis, which limits the ability to identify the key drivers of variation in efficiency across settings.

https://doi.org/10.69588/mjer.v30i19.5654

Future research should aim to improve the comparability and policy relevance of DEA-based efficiency studies in health systems. First, standardized definitions of DMUs are needed, particularly for primary care facilities, district health systems, and long-term care institutions. These definitions should consider organizational structure, service scope, ownership, and regulatory context. Second, future DEA models should more routinely incorporate quality indicators, such as patient-reported outcomes, clinical process measures, safety indicators, or composite quality indices. This would help ensure that efficiency is not interpreted solely as producing more services with fewer resources, but also as delivering appropriate and high-quality care. Third, future studies should apply contextual adjustment methods, such as two-stage DEA or meta-frontier analysis, to distinguish managerial inefficiency from environmental or structural constraints. These approaches should be accompanied by appropriate statistical inference, as recommended by Simar and Wilson (2007) and O'Donnell, Rao, and Battese (2008). Fourth, longitudinal approaches, including Malmquist productivity indices or dynamic DEA, should be used to examine changes in efficiency over time and to decompose productivity change into technical efficiency change, technological change, and scale efficiency change (Färe, Grosskopf, Norris, & Zhang, 1994). Finally, more implementation-oriented research is needed to evaluate whether DEA-based performance assessment can meaningfully improve decision-making, resource allocation, service delivery, quality of care, and equity. Such studies would help move DEA from a purely analytical tool toward a practical instrument for health system strengthening.

V. CONCLUSION AND RECOMMENDATIONS

This systematic review synthesized DEA applications in primary care, district health systems, and long-term care, revealing substantial efficiency gaps but also significant methodological heterogeneity. Pooled mean efficiency scores ranged from 0.458 to 0.608, indicating potential for output expansion of 64% to 118% if DMUs operated at the efficiency frontier. However, heterogeneity in DMU definitions, reliance on volume-based outputs, and limited quality adjustment constrain direct policy inference.

Methodological recommendations:

1. Stratify DMUs by organizational type, service scope, and regulatory environment to ensure comparability.
2. Incorporate quality indicators into DEA models using quality-adjusted outputs, undesirable outputs, or two-stage quality adjustment.
3. Explicitly justify DEA model specification (CRS/VRS, orientation) based on institutional context.
4. Apply two-stage DEA or meta-frontier analysis to separate managerial inefficiency from environmental constraints.
5. Report bootstrap confidence intervals to quantify uncertainty.

Policy recommendations:

1. Interpret DEA efficiency scores as context-sensitive benchmarks, not absolute standards.
2. Link efficiency measurement to capacity-building, process improvement, and management support rather than punitive budget cuts.

https://doi.org/10.69588/mjer.v30i19.5654

3. Triangulate DEA results with qualitative evidence, stakeholder consultation, and equity analysis before implementing resource allocation decisions.
4. Invest in health information systems and data quality to enable routine, reliable efficiency measurement.

For resource-constrained public health systems, including Mongolia, DEA offers a practical tool for efficiency measurement and resource allocation, but its effective application requires methodological rigor, context-sensitive interpretation, and sustained investment in data infrastructure and analytical capacity.

REFERENCES

AbouZahr, C., & Boerma, T. (2005). Health information systems: The foundations of public health. *Bulletin of the World Health Organization, 83*(8), 578–583. https://www.who.int/bulletin/volumes/83/8/578.pdf

Afonso, A., & St. Aubyn, M. (2006). *Relative efficiency of health provision: A DEA approach with non-discretionary inputs* (ISEG Economics Working Paper No. 33/2006/DE/UECE). https://doi.org/10.2139/ssrn.952629

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Banker, R. D., Conrad, R. F., & Strauss, R. P. (1986). A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital production. *Management Science*, *32*(1), 30–44. https://doi.org/10.1287/mnsc.32.1.30

Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration, 84*(3), 517–538. https://doi.org/10.1111/j.1467-9299.2006.00600.x

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. Springer. https://doi.org/10.1007/978-1-4419-7961-2

Cantor, V. J. M., & Poh, K. L. (2018). Integrated analysis of healthcare efficiency: A systematic review. *Journal of Medical Systems*, *42*, Article 8. https://doi.org/10.1007/s10916-017-0848-7

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research, 2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Cheng, Z., Tao, H., Cai, M., Lin, H., Lin, X., Shu, Q., & Zhang, R. N. (2015). Technical efficiency and productivity of Chinese county hospitals: An exploratory study in Henan province, China. *BMJ Open, 5*(9), Article e007267. https://doi.org/10.1136/bmjopen-2014-007267

Chilingerian, J. A., & Sherman, H. D. (2004). Health care applications: From hospitals to physicians, from productive efficiency to quality frontiers. In W. W. Cooper, L. M. Seiford, & J. Zhu (Eds.), *Handbook on data envelopment analysis* (pp. 481–537). Springer. https://doi.org/10.1007/1-4020-7798-X_16

Clement, J. P., Valdmanis, V. G., Bazzoli, G. J., Zhao, M., & Chukmaitov, A. (2008). Is more better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion. *Health Care Management Science, 11*(1), 67–77. https://doi.org/10.1007/s10729-007-9025-8

https://doi.org/10.69588/mjer.v30i19.5654

Coelli, T., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). Springer. https://doi.org/10.1007/b136381

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-45283-8

Cylus, J., Papanicolas, I., & Smith, P. C. (Eds.). (2016). *Health system efficiency: How to make measurement matter for policy and management*. WHO Regional Office for Europe. https://www.euro.who.int/en/publications/abstracts/health-system-efficiency-2016

Donabedian, A. (1988). The quality of care: How can it be assessed? *JAMA, 260*(12), 1743–1748. https://doi.org/10.1001/jama.1988.03410120089033

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford University Press.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research, 132*(2), 245–259. https://doi.org/10.1016/S0377-2217(00)00149-1

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General), 120*(3), 253–290. https://doi.org/10.2307/2343100

Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review, 84*(1), 66–83. https://www.jstor.org/stable/2117971

Giuffrida, A. (1999). Productivity and efficiency changes in primary care: A Malmquist index approach. *Health Care Management Science, 2*(1), 11–26. https://doi.org/10.1023/A:1019027009061

Giuffrida, A., & Gravelle, H. (2001). Measuring performance in primary care: Econometric analysis and DEA. *Applied Economics, 33*(2), 163–175. https://doi.org/10.1080/00036840122522

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. *Health Economics, 17*(10), 1107–1128. https://doi.org/10.1002/hec.1391

Hollingsworth, B., & Peacock, S. J. (2008). *Efficiency measurement in health and health care*. Routledge. https://doi.org/10.4324/9780203486566

Hussey, P. S., de Vries, H., Romley, J., Wang, M. C., Chen, S. S., Shekelle, P. G., & McGlynn, E. A. (2009). A systematic review of health care efficiency measures. *Health Services Research, 44*(3), 784–805. https://doi.org/10.1111/j.1475-6773.2008.00942.x

Jacobs, R., Smith, P. C., & Street, A. (2006). *Measuring efficiency in health care: Analytic techniques and health policy*. Cambridge University Press. https://doi.org/10.1017/CBO9780511617492

Joanna Briggs Institute. (2016). *Critical appraisal checklist for economic evaluations*. JBI. https://jbi.global/critical-appraisal-tools

Kane, R. L., Shamliyan, T., Mueller, C., Duval, S., & Wilt, T. J. (2007). The association of registered nurse staffing levels and patient outcomes: Systematic review and meta-analysis. *Medical Care, 45*(12), 1195–1204. https://doi.org/10.1097/MLR.0b013e3181468ca3

Kohl, S., Schoenfelder, J., Fügener, A., & Brunner, J. O. (2019). The use of data envelopment analysis (DEA) in healthcare with a focus on hospitals. *Health Care Management Science, 22*(2), 245–286. https://doi.org/10.1007/s10729-018-9436-8

Kringos, D. S., Boerma, W. G. W., Hutchinson, A., van der Zee, J., & Groenewegen, P. P. (2010). The breadth of primary care: A systematic literature review of its core dimensions. *BMC Health Services Research, 10*, Article 65. https://doi.org/10.1186/1472-6963-10-65

Kruk, M. E., Gage, A. D., Arsenault, C., Jordan, K., Leslie, H. H., Roder-DeWan, S., Adeyi, O., Barker, P., Daelmans, B., Doubova, S. V., English, M., García-Elorrio, E., Guanais, F., Gureje, O., Hirschhorn, L. R., Jiang, L., Kelley, E., Lemango, E. T., Liljestrand, J., ... Pate, M. (2018). High-quality health systems in the Sustainable Development Goals era: Time for a revolution. *The Lancet Global Health, 6*(11), e1196–e1252. https://doi.org/10.1016/S2214-109X(18)30386-3

McGarvey, R. G., Thorsen, A., Thorsen, M. L., & Hozain, S. (2019). Measuring efficiency of community health centers: A multi-model approach considering quality of care and heterogeneous operating environments. *Health Care Management Science, 22*(3), 489–504. https://doi.org/10.1007/s10729-018-9455-5

Moran, V., Suhrcke, M., & Nolte, E. (2023). Exploring the association between primary care efficiency and health system characteristics across European countries: A two-stage data envelopment analysis. *BMC Health Services Research, 23*, Article 1348. https://doi.org/10.1186/s12913-023-10369-y

Nayar, P., & Ozcan, Y. A. (2008). Data envelopment analysis comparison of hospital efficiency and quality. *Journal of Medical Systems, 32*(3), 193–199. https://doi.org/10.1007/s10916-007-9122-8

O'Donnell, C. J., Rao, D. S. P., & Battese, G. E. (2008). Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics, 34*(2), 231–255. https://doi.org/10.1007/s00181-007-0119-4

OECD. (2020). *Who cares? Attracting and retaining care workers for the elderly*. OECD Publishing. https://doi.org/10.1787/92c0ef68-en

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pelone, F., Kringos, D. S., Spreeuwenberg, P., De Belvis, A. G., & Groenewegen, P. P. (2015). Primary care efficiency measurement using data envelopment analysis: A systematic review. *Journal of Medical Systems, 39*, Article 156. https://doi.org/10.1007/s10916-014-0156-4

Purohit, B. (2016). Healthcare sector efficiency in Gujarat (India): An exploratory study using data envelopment analysis. *Health Sciences, 5*(5), 55–64. https://doi.org/10.4081/hls.2016.5525

Rahimi, H., Bahmaei, J., Shojaei, P., Kavosi, Z., & Khavasi, M. (2021). *Performance evaluation of the district's primary health care system: A case study of southeastern Iran*. Research Square. https://doi.org/10.21203/rs.3.rs-393521/v1

Ramirez-Valdivia, M. T., Maturana-Ross, S., & Salvo-Garrido, S. (2015). Measuring the efficiency of Chilean primary healthcare centres. *International Journal of Engineering Business Management, 7*, Article 31. https://doi.org/10.5772/60839

Renner, A., & Kirigia, J. M. (2005). Technical efficiency of peripheral health units in Pujehun District of Sierra Leone: A DEA application. *African Health Economics Monitor, 6*(1), 2–7.

Shao, L., Yu, X., & Feng, C. (2021). A SBM-DEA based performance evaluation and optimization for social organizations participating in community and home-based elderly care services. *PLOS ONE, 16*(3), Article e0248474. https://doi.org/10.1371/journal.pone.0248474

Sherman, H. D. (1984). Hospital efficiency measurement and evaluation: Empirical test of a new technique. *Medical Care, 22*(10), 922–938. https://doi.org/10.1097/00005650-198410000-00005

Shimshak, D. G., Lenard, M. L., & Klimberg, R. K. (2009). Incorporating quality into data envelopment analysis of nursing home performance: A case study. *Omega, 37*(3), 672–685. https://doi.org/10.1016/j.omega.2008.05.004

Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science, 44*(1), 49–61. https://doi.org/10.1287/mnsc.44.1.49

Simar, L., & Wilson, P. W. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis, 13*(1), 49–78. https://doi.org/10.1023/A:1007864806704

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics, 136*(1), 31–64. https://doi.org/10.1016/j.jeconom.2005.07.009

Smith, P. C., & Street, A. (2005). Measuring the efficiency of public services: The limits of analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168*(2), 401–417. https://doi.org/10.1111/j.1467-985X.2005.00355.x

Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *The Milbank Quarterly, 83*(3), 457–502. https://doi.org/10.1111/j.1468-0009.2005.00409.x

Su, W., Du, L., Fan, Y., & Wang, P. (2023). Evaluating the efficiency of primary health care institutions in China: An improved three-stage data envelopment analysis approach. *BMC Health Services Research, 23*, Article 1035. https://doi.org/10.1186/s12913-023-09979-3

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130*(3), 498–509. https://doi.org/10.1016/S0377-2217(99)00407-5

Tone, K., & Tsutsui, M. (2010). Dynamic DEA: A slacks-based measure approach. *Omega, 38*(3–4), 145–156. https://doi.org/10.1016/j.omega.2009.07.003

World Health Organization. (2008). *Integrated health services: What and why?* WHO. https://www.who.int/publications/i/item/WHO-HIS-SDS-2018.50

World Health Organization. (2010). *The world health report: Health systems financing: The path to universal coverage*. WHO. https://www.who.int/publications/i/item/9789241564021

Worthington, A. C. (2004). Frontier efficiency measurement in health care: A review of empirical techniques and selected applications. *Medical Care Research and Review, 61*(2), 135–170. https://doi.org/10.1177/1077558704263796

Yitbarek, K., Abraham, G., Adamu, A., Tsega, G., Berhane, M., Hurlburt, S., Mann, C., Woldie, M., & Lindelow, M. (2021). Significant inefficiency in running community health systems:

https://doi.org/10.69588/mjer.v30i19.5654

The case of health posts in Southwest Ethiopia. *PLOS ONE*, *16*(2), Article e0246559. https://doi.org/10.1371/journal.pone.0246559

Zakowska, I., Godycki-Cwirko, M., & Czerw, A. (2020). Data envelopment analysis applications in primary health care: A systematic review. *Family Practice, 37*(2), 147–153. https://doi.org/10.1093/fampra/cmz057

Zavras, A. I., Tsakos, G., Economou, C., & Kyriopoulos, J. (2002). Using DEA to evaluate efficiency and formulate policy within a Greek national primary health care network. *Journal of Medical Systems*, *26*(4), 285–292. https://doi.org/10.1023/A:1015860318972