Health care quality in nonparametric efficiency studies: a review

Margit Sommersguter-Reichmann

Accepted: 23 August 2021 / Published online: 30 September 2021
© The Author(s) 2021

Abstract
Scientists are increasingly considering quality in nonparametric frontier efficiency studies in health care. There are many ways to include quality in efficiency analyses. These approaches differ, among other things, in the underlying assumptions about the influence of quality on the attainable efficiency frontier and the distribution of inefficiency scores. The aim is to provide an overview of how scholars have taken quality into account in nonparametric frontier efficiency studies and, at the same time, to address the underlying assumptions on the relationship between efficiency and quality. To this end, we categorized empirical efficiency studies according to the methodological approaches and quality dimensions and collected the quality indicators used. We performed a Web of Science search for studies published in journals covered by the Science Citation Index Expanded, the Social Sciences Citation Index, and the Emerging Sources Citation Index between 1980 and 2020. Of the 126 studies covered in this review, 78 are one-stage studies that incorporate quality directly into the efficiency model and thus assume that quality impacts the attainable efficiency frontier. Forty-four articles are two-stage studies that consider quality in the first and the second stage or the second stage only. Four studies do not assume a priori a specific association between efficiency and quality. Instead, they test for this relationship empirically. Outcome quality is by far the most frequently incorporated quality dimension. While most studies consider structural quality as an environmental variable in the second stage, they include outcome quality predominantly directly in the efficiency model. Process quality is less common.

Keywords Review · Nonparametric efficiency · Data envelopment analysis · Health care · Quality
Abbreviations

ACE          Angiotensin-converting enzyme
ADL          Activities of daily living
ALOS         Average length of stay
AMI          Acute myocardial infarction
ARB          Angiotensin receptor blocker
CHF          Congestive heart failure
CMI          Case-mix index
COPD         Chronic obstructive pulmonary disease
CVD          Cardiovascular disease
CVRM         Cardiovascular risk management
ED           Emergency department
ER           Emergency room
EURO-QOL 5D5L European quality of life 5 dimensions 5 level version
FTE          Full-time equivalent
GH           Gastrointestinal hemorrhage
HA           Heart attack
HbA1c        Glycated hemoglobin
HF           Heart failure
HIV          Human immunodeficiency virus
HLY          Healthy life year
LDL cholesterol Low-density lipoprotein cholesterol
LVS          Left ventricular systolic function
MDS          Minimum data set
OCP          On-call physician
OR           Operating room
PN           Pneumonia
PYLG         Potential years of life gained
PYLL         Potential years of life lost
RN           Registered nurse
TQM          Total quality management

1 Introduction

Differences in the cost of providing comparable health care services have led to efforts to identify efficient service providers. Efficiency in service provision can be measured by the ratio of actual to best practice productivity, which, in turn, corresponds to the ratio of output (e.g., case-mix adjusted inpatient days) to input (e.g., health care personnel). Many studies have used nonparametric frontier techniques, such as Data Envelopment Analysis (DEA), to assess the efficiency of the health care sector (see, e.g., Hollingsworth et al. 1999; O’Neill et al. 2008; Kohl et al. 2018). Nonparametric frontier methods can simultaneously consider multiple inputs and outputs.
without knowing their functional relationship and the factor and product prices. Service providers are rated efficient if they produce a given output with minimum input or the maximum output for a given input, compared with other service providers.

However, the health economics literature increasingly recognizes that in addition to concentrating on quantity input and quantity output, health care quality requires due consideration. According to Donabedian (1980), health care quality consists of structural, process, and outcome quality. Portela et al. (2016, p. 727) outline that the structural quality represents the characteristics of the health care setting, the process quality refers to the various actions undertaken by health care staff in providing care, and the outcome quality reflects the effects of health care services on the patient’s condition. While there are other classifications of quality, we used the Donabedian framework to apply a definition of quality that was valid through the overall review period.

Incorporating quality in efficiency studies is a complex task. As early as 1970, Newhouse (1970) elaborated that quantity and quality are goods that compete for resources. A service provider has to compromise between increasing the quantity and increasing the quality given scarce resources. Many studies investigated the theory of the efficiency-quality trade-off in the health care sector. In their review, Hussey et al. (2013) examined the cost-quality trade-off with the following, almost evenly distributed results: around one-third of studies (34%) identify a positive or mixed-positive relationship, slightly less than one-third (30%) reveal a negative or mixed-negative relationship, and somewhat more than one-third (36%) show a lack of, an ambiguous, or a mixed association. The results in the nonparametric efficiency literature are also diverse. While Kooreman (1994) finds evidence of a trade-off between efficiency and quality in assessing the efficiency of Dutch nursing homes, Laine et al. (2005) identify mixed results regarding the association between different quality indicators and the efficiency of (hospital and nursing care) wards providing long-term care in Finland. Pross et al. (2018) reveal a trade-off between quality and efficiency in providing stroke care in German hospitals. Khushalani and Ozcan (2017) do not find evidence of any efficiency-quality trade-off when investigating the efficiency of US hospitals.

The mixed results regarding the relationship between efficiency and quality in nonparametric studies are not surprising, as the studies differ significantly in several factors. First, the units of analysis (decision-making units, DMUs) differ in terms of type (e.g., hospitals, nursing homes) and aggregation level (e.g., facility level, ward level, patient-level). Second, the studies cover different quality dimensions and use varying quality indicators. Third, the DMUs operate in different environments, each of which can have other effects on efficiency. Fourth, the model specifications, most of which assume a priori a specific efficiency-quality association and are tailored to particular research questions, vary. Study results are therefore less likely to be generally valid. Instead, they require interpretation in the respective context.

The paper, however, aims to provide an overview of how scholars have integrated quality into efficiency studies. In particular, the aim is to classify nonparametric efficiency studies according to model specifications and quality dimensions. For that purpose, we slightly modified the classifications of Bădin et al. (2014) and Varabyova

---

1 For a summary, see Busse et al. (2019, p. 5).
et al. (2017) to particularly pay attention to the model-inherent assumptions about how quality affects efficiency. In addition, the paper gives a comprehensive overview of the quality indicators considered in nonparametric efficiency studies.

The remainder of this paper is as follows. Section 2 describes the data sources, search strategy, and study selection. It also presents the classification scheme and discusses the quality dimensions. Section 3 describes the results. Section 4 discusses challenges when considering structure, process, and outcome quality in efficiency studies. The paper concludes with a summary and a discussion of limitations in Sect. 5.

2 Methods

2.1 Data sources, search strategy, and study selection

The selection of the relevant publications was as follows. We performed a Web of Science (WoS) keyword search to identify nonparametric frontier efficiency studies related to health care. We searched for studies published between 1980 and 2020 covered by the Science Citation Index Expanded, the Social Sciences Citation Index, and the Emerging Sources Citation Index. We then examined the studies’ keywords and abstracts for nonparametric frontier techniques. We recorded the model specification, the type of DMU, the quality dimensions, and the quality indicators for those studies that applied nonparametric methods. Finally, we analyzed cross-references to studies not covered by the WoS search for possible inclusion in the review.

2.2 Classification of model specifications

We used a classification that accounts for a priori assumptions regarding the relationship between efficiency and quality. To that end, we slightly modified the categorization proposed by Bădăin et al. (2014) and Varabyova et al. (2017). First, we distinguished between Category 1 (CAT1) and Category 2 (CAT2) studies. In contrast to CAT2, CAT1 comprises studies that make a priori assumptions about the association between quality and efficiency. CAT1 consists of one- and two-stage approaches referred to as Class 1 and Class 2 studies.

Class 1 studies incorporate quality directly into the efficiency model, assuming that quality affects the attainable efficiency frontier. In contrast to the conventional classification in the literature, we only classified studies that consider quality in the second stage as Class 2 approaches because the focus is on how quality is incorporated. Put differently, when a paper accounts for other than quality-related factors in the second stage, we classified the article as a Class 1 study. Accounting for quality in the second stage corresponds to assuming that quality affects the inefficiency distribution and, thus, the separability between the input–output-space and the quality-space to hold (Simar and Wilson 2007).

Class 1 includes the five subcategories (SubCATs) 11 to 15 (Table 1):

2 Strictly speaking, one might also test whether the separability condition holds; see Daraio et al. (2010).
### Table 1: Classification scheme

| Class | Stage | Subcategories | Description | Characteristics |
|-------|-------|---------------|-------------|-----------------|
| 1, 2  | 1     | 11, 21        | x, y, q     | Account for quality directly in the efficiency model, i.e., assume that quality and efficiency are interrelated performance dimensions. |
|       |       | 12, 22        | x, y, AND x, q | Evaluate efficiency and quality separately, i.e., assume that quality and efficiency represent distinct performance dimensions. |
|       |       | 13, 23        | x, q        | Consider quality only; this corresponds to assessing the clinical efficiency, that is, the effectiveness, rather than managerial efficiency. |
|       |       | 14, 24        | x, y, AND x, y, q | Evaluate the efficiency once with, once without quality. The primary purpose is to examine the effect of quality on efficiency. |
|       |       | 15, 25        | x, y (q), AND y (q) → q | Analyze the input–output–outcome relationship. |
| 2     | 2     | 20–25         | eff (q) → q (eff) | Evaluate whether differences in quality explain efficiency (effectiveness) differences or vice versa. |

x, input; y, output; q, quality; eff, efficiency (effectiveness)

- **SubCAT 11** studies incorporate quality in the efficiency model, assuming that efficiency (measured in input and output quantities) and quality are related performance dimensions.
- **SubCAT 12** studies use separate models for efficiency and quality. This approach understands efficiency and quality as distinct performance dimensions, resulting in an efficiency (also called managerial) model and an effectiveness (also called clinical) model (see, e.g., Portela et al. 2016).
- **SubCAT 13** covers effectiveness studies that only incorporate quality outputs.
- **SubCAT 14** studies examine the effect of quality on efficiency by using separate models. In contrast to SubCAT 12, however, they account for quality and efficiency in one specification while omitting quality in the other.
- **SubCAT 15** comprises Network DEA (NDEA) studies that overcome the black-box assumption regarding transforming input into outcome (final output). NDEA allows decomposing the overall service production process into two or more corresponding subprocesses. In the health care context, NDEA provides the opportunity to model the input–output–outcome relationship.
Class 2 studies comprise the six SubCATs 20 to 25 (see Table 1 for details). SubCAT 20 has no equivalent in Class 1 since these studies do not take quality into account directly in the efficiency model (i.e., the first stage), but only in the second stage. The model specifications in the first stage in SubCATs 21 to 25 correspond to those in SubCATs 11 to 15. Class 2 papers examine the relationship between efficiency (or effectiveness) and (further) quality indicators in the second stage, mainly using regression analysis. The inclusion of quality in a second stage corresponds to the assumption that quality affects the distribution of inefficiencies; that is, the separability condition holds.

In contrast to CAT1, CAT2 studies do not a priori assume a specific association between efficiency and quality. Instead, these studies rely on conditional efficiency so that "the data themselves tell if and how they (the data or quality in our case; author's note) affect the performance" (Daraio and Simar 2007b, p. 14). Bădin et al. (2012) illustrate that regressing the ratio of conditional to unconditional efficiency on quality identifies an influence of quality on the efficiency frontier while regressing conditional efficiency on quality reveals any impact of quality on the inefficiency distribution.

### 2.3 Quality dimensions and corresponding indicators

According to Donabedian (1980), structural quality represents the characteristics of the health care setting, represented by indicators such as appropriate equipment and staff mix. Structural quality is patient-invariant. The possibilities of including structural quality in efficiency studies are pretty heterogeneous. It is possible to incorporate structural quality as a nondiscretionary input (Banker and Morey 1986), desirable output, or a quality-adjusted output directly in the efficiency model.

Clinical guidelines and good clinical practice are often the basis for process quality indicators (Quentin et al. 2019, p. 39), which vary across patient characteristics. Efficiency studies usually incorporate process indicators as outputs. However, medical care consists of several successive phases. Therefore, process indicators can be considered output (referred to as intermediate output or throughput) in one phase and input in the other. Process quality can also be associated with managerial aspects of medical care.

For patients, outcome quality is probably the most crucial quality dimension. Subjective and objective indicators can depict outcome quality. While subjective indicators reflect the opinions and assessments of patients, objective indicators rely on measurable data such as mortality rates, readmission rates, and complication rates. Outcome quality is the final output of providing medical care and also depends on patient characteristics.

Some studies incorporate outcome quality using desirable, others using undesirable indicators. For desirable indicators, such as the patients’ willingness to recommend a hospital, the higher the quality, the better. For undesirable indicators such as the mortality rate, there is an inverse relationship between quality and efficiency. Incorporating undesirable indicators into a nonconditional nonparametric efficiency model

3 Occasionally, process indicators are also measured as undesirable indicators; see, e.g., Giménez et al. (2019).
requires special attention. The literature broadly distinguishes between approaches with and without data transformations (see, for example, Liu et al. (2010) and Halkos and Petrour (2019) for a summary of the range and characteristics of the different approaches).

Models with data transformations include the linear additive transformation suggested by Ali and Seiford (1990), Pastor (1996), and Scheel (2001), and the nonlinear multiplicative transformation proposed by Lovell et al. (1987) and Golany and Roll (1989). Models without data transformation involve treating undesirable outputs as desirable inputs (Liu and Sharp 1999), assuming undesirable outputs to be weakly disposable (e.g., that an increase in quality comes at the expense of quantity outputs), and using sophisticated efficiency metrics (e.g., directional distance functions (DDFs), slack-based models (SBMs), and hyperbolic measures).

### 3 Results

The WOS keyword search, which covered the period from 1980 until 2020, yielded 504 studies. After analyzing keywords and abstracts, 119 of 185 nonparametric efficiency studies included at least one quality indicator. Reference mining and ad hoc searches returned seven additional articles so that the review finally covers 126 studies. Table 2 provides the complete list of reviewed studies, including their categorization. We included the categorization because, in some cases, multiple classifications would have been possible. The same applies to the assignment of the indicators to the corresponding quality dimensions. In this regard, we refer the interested reader to the Appendix.

#### 3.1 Studies by categories, classes, and subcategories

Table 3 shows the categorization of the studies. 122 studies (97%) belong to CAT1, leaving four CAT2 studies. This finding is not surprising. After all, conditional efficiency approaches are relatively new methods. Within CAT1, 78 studies are one-stage approaches (Class 1). SubCAT 11 (27 papers) dominates in this class. Effectiveness studies (SubCAT 13: 23 papers) build the second most common subcategory, followed by studies that examine the effect of quality on efficiency by adding at least one quality indicator in a separate specification (SubCAT 14: 17 papers). Studies that understand efficiency and quality as distinct dimensions (SubCAT 12: 5 papers), and NDEA studies (SubCAT 15: 6 papers), are less common.

SubCAT 20 studies that only consider quality in Stage 2 dominate in Class 2 (19 papers). In second place are SubCAT 21 studies (12 papers). Effectiveness studies (SubCAT 23: 8 papers) are also quite common, as opposed to studies in SubCATs 22 (1 paper), 24 (3 papers), and 25 (1 paper).

---

4 The last search query was performed on November 13, 2020.
### Table 2: List of studies included in the review

| SubCAT | References | DMU | Orientation | Undesirable outputs |
|--------|------------|-----|-------------|---------------------|
| 12     | Kočišová et al. (2020) | COUN | IO          | Data transformation |
| 12     | Lo Storto and Goncharuk (2017) | COUN | IO/OO       | Input               |
| 13     | Adang and Born (2007) | COUN | OO          | Data transformation |
| 13     | Althin et al. (2019) | COUN | IO          | Data transformation |
| 13     | Kozuń-Cieślak (2020) | COUN | NO          | Data transformation |
| 13     | Ortega et al. (2017) | COUN | IO          | Data transformation |
| 13     | Santos et al. (2012) | COUN | OO          |                     |
| 15     | Kim et al. (2020) | COUN | OO          | NDEA                |
| 23     | Puertas et al. (2020) | COUN | OO          | Data transformation |
| 11     | Arrieta and Guillén (2017) | HC   | OO          | WD                  |
| 11     | Azadeh et al. (2016) | HC   | IO          | Output              |
| 11     | Bilsel and Davutyan (2014) | HC   | NO          | WD and DDF          |
| 11     | Chang et al. (2011) | HC   | IO          | Data transformation |
| 11     | Chen et al. (2016) | HC   | IO          | Output              |
| 11     | Chen et al. (2019) | HC   | IO          | Output              |
| 11     | Du et al. (2014) | HC   | NO          |                     |
| 11     | Giménez et al. (2019) | HC   | OO          | DDF                 |
| 11     | Guo et al. (2017) | HC   | NO          | Input               |
| 11     | Hu et al. (2012) | HC   | OO          | Data transformation |
| 11     | Kalis and Stracova (2019) | HC   | IO          | Nondiscretionary factor |
| 11     | Matranga and Sapienza (2015) | HC   | OO          | WD                  |
| 11     | Miller et al. (2017) | HC   | NO          |                     |
| 11     | Morey et al. (1995) | HC   | IO          | Nondiscretionary factor |
| 11     | Nedelea and Fannin (2017) | HC   | IO          |                     |
| 11     | Park et al. (2018) | HC   | IO          |                     |
| 11     | Sinuany-Stern et al. (2016) | HC   | OO/NO       | Data transformation |
| 11     | Thanassoulis et al. (2016) | HC   | NO          |                     |
| 11     | Wu et al. (2013) | HC   | OO          | WD and DDF          |
| 13     | Chang and Lan (2010) | HC   |             | Data transformation |
| 13     | Dismuke and Sena (2001) | HC   | OO          | WD and DDF          |
| SubCAT | References                                         | DMU | Orientation | Undesirable outputs |
|--------|---------------------------------------------------|-----|-------------|---------------------|
| 13     | Kholghabad et al. (2019)                          | HC  | IO/OO       | Input               |
| 13     | Williams et al. (2016)                            | HC  | IO          | Output              |
| 14     | Ferrier and Trivitt (2013)                         | HC  | OO/IO       |                     |
| 14     | Karagiannis and Velentzas (2012)                  | HC  | OO          |                     |
| 14     | Navarro-Espigares and Torres (2011)               | HC  | IO          | Output              |
| 14     | Nayar and Ozcan (2008)                            | HC  | IO          |                     |
| 14     | Prior (2006)                                      | HC  | OO          | Input and WD (output) |
| 14     | Sola and Prior (2001)                             | HC  | IO          | Output              |
| 14     | Tiemann and Schreyögg (2009)                      | HC  | IO          | Data transformation |
| 14     | van Ineveld et al. (2016)                         | HC  | IO          | Data transformation |
| 14     | Yang and Zeng (2014)                              | HC  | OO          | Data transformation |
| 15     | Cohen-Kadosh and Sinuany-Stern (2020)             | HC  | OO          | Data transformation and NDEA |
| 20     | Chu et al. (2002)                                 | HC  | OO          |                     |
| 20     | Ferreira et al. (2020)                            | HC  | NO          | Conditional DDF<sup>d</sup> |
| 20     | Ferrier and Valdmanis (1996)                      | HC  | IO          |                     |
| 20     | Gautam et al. (2013)                              | HC  | IO          |                     |
| 20     | Gholami et al. (2015)                             | HC  | IO          |                     |
| 20     | Gok and Sezen (2013)                              | HC  | IO          |                     |
| 20     | Harrison and Coppola (2007)                       | HC  |             |                     |
| 20     | Huerta et al. (2011)                              | HC  |             |                     |
| 20     | Kittelsen et al. (2015)                           | HC  | IO          |                     |
| 20     | Lindlbauer et al. (2016)                          | HC  | IO          |                     |
| 20     | Mujasi et al. (2016)                              | HC  | OO          |                     |
| 20     | Pai et al. (2019)                                 | HC  | IO          |                     |
| 21     | Barpanda and Sreekumar (2020)                     | HC  | IO          | Output              |
| 21     | Clement et al. (2008)                             | HC  | OO          | WD                  |
| 21     | Dutta et al. (2014)                               | HC  | OO          |                     |
| 21     | Leleu et al. (2018)                               | HC  | OO          | DDF                 |
| 21     | Lin et al. (2019)                                 | HC  | OO          | Data transformation |
| 21     | Morey et al. (1992)                               | HC  | IO          | Nondiscretionary factor |
| 21     | Mutter et al. (2010)                              | HC  | OO/IO       | WD                  |
| 21     | Nayar et al. (2013)                               | HC  | IO          | Data transformation |
Table 2 (continued)

| SubCAT | References | DMU a | Orientation b | Undesirable outputs c |
|--------|------------|-------|---------------|-----------------------|
|        |            |       |               |                       |
| 21     | Valdmanis et al. (2008) | HC    | OO/OO         | WD                    |
| 21     | Wang et al. (2018)       | HC    | OO            | Data transformation   |
| 22     | Roth et al. (2019)       | HC    | IO            | Data transformation   |
| 24     | Chiu et al. (2008)       | HC    | IO            |                       |
| 24     | Li (2015)                | HC    | IO            | Input                |
| 25     | Khushalani and Ozcan (2017) | HC   | NO            | NDEA                 |
| 30     | Ferreira and Marques (2019) | HC    | NO            | Conditional directional order-α |
| 30     | Varabyova et al. (2017)  | HC    | OO            | Conditional order-m   |
| 20     | Laine et al. (2005)      | HC plus NC/AC | OO          |
| 14     | Obure et al. (2016)      | HC plus PHC | OO          |
| 11     | Campoli et al. (2020)    | HR    | OO            | Data transformation   |
| 11     | Giuffrida and Gravelle (2001) | HR | IO            | Nondiscretionary factor |
| 11     | Karsak and Karadayi (2017) | HR | IO            |                       |
| 11     | Thanassoulis et al. (1995) | HR    | OO/OO         | Nondiscretionary factor, data transformation |
| 11     | Zheng et al. (2019)      | HR    | OO            | Input                |
| 13     | Carrillo and Jorge (2017) | HR    | IO            | Data transformation   |
| 13     | Ngobeni et al. (2020)    | HR    | IO            | Output               |
| 13     | Pelone et al. (2012)     | HR    | IO            | Output               |
| 13     | Salinas-Jiménez and Smith (1996) | HR | IO            |                       |
| 13     | Serván-Mori et al. (2018) | HR | OO            |                       |
| 13     | Takundwa et al. (2017)   | HR    | IO            | Data transformation   |
| 14     | Du (2018)                | HR    | NO            |                       |
| 14     | Polisena et al. (2010)   | HR    | OO            | Output               |
| 23     | Allin et al. (2016)      | HR    | OO            | Data transformation   |
| 23     | Amado and Dyson (2009)   | HR    | IO/OO         |                       |
| 23     | Campanella et al. (2017) | HR    | IO            | Output               |
| 23     | De et al. (2012)         | HR    | OO            | Data transformation   |
| 23     | Serván-Mori et al. (2020) | HR | OO            |                       |
| 30     | Ferreira et al. (2013)   | HR    | IO/OO/NO      | Conditional order-m, DEA |
| 30     | Gearhart and Michieka (2020) | HR | OO            | Conditional order-m   |
| SubCAT | References                                      | DMU | Orientation | Undesirable outputs                           |
|--------|------------------------------------------------|-----|-------------|-----------------------------------------------|
| 12     | Shimshak and Lenard (2007)                      | NC/AC | IO          | Data transformation                           |
| 13     | Bott et al. (2007)                              | NC/AC | IO          | Data transformation                           |
| 13     | Min et al. (2016)                               | NC/AC | IO          | Data transformation                           |
| 13     | Shen et al. (2020)                              | NC/AC | IO          | Data transformation                           |
| 13     | Shwartz et al. (2016)                           | NC/AC | IO          | Input                                         |
| 14     | Duffy et al. (2006)                             | NC/AC | IO          | Data transformation                           |
| 14     | Kleinsorge and Karney (1992)                    | NC/AC | OO          | Input and data transformation                 |
| 15     | Avkiran and McCrystal (2014)                    | NC/AC | IO          | NDEA                                          |
| 20     | Kooreman (1994)                                 | NC/AC | IO          |                                               |
| 20     | Lee et al. (2009)                               | NC/AC | IO          |                                               |
| 20     | Lin et al. (2017)                               | NC/AC | IO          |                                               |
| 20     | Ni Luasa et al. (2018)                          | NC/AC | IO          |                                               |
| 21     | Garavaglia et al. (2011)                        | NC/AC | IO          |                                               |
| 21     | Zhang et al. (2019)                             | NC/AC | IO          | Output                                        |
| 23     | Wichmann et al. (2018)                          | NC/AC | OO          |                                               |
| 24     | Zhang et al. (2008)                             | NC/AC | IO          | Data transformation                           |
| 13     | Villamil et al. (2017)                          | OTHER (INS) | OO | Hyperbolic efficiency                        |
| 20     | Bryce et al. (2000)                             | OTHER (INS) | IO |                                               |
| 15     | Ghafari Someh et al. (2020)                     | OTHER (LAB) | IO | NDEA                                          |
| 20     | Lamovšek and Klun (2020)                        | OTHER (LAB) | IO |                                               |
| 15     | Borg et al. (2019)                              | OTHER (PAT) | OO |                                               |
| 13     | Dowd et al. (2014)                              | OTHER (PHY) | IO | Data transformation                           |
| 13     | Fiallos et al. (2017)                           | OTHER (PHY) | NO | Data transformation                           |
| 11     | Aduda et al. (2015)                             | PHC | OO          |                                               |
| 11     | Milliken et al. (2011)                          | PHC | OO          |                                               |
| 11     | Ruiz-Rodriguez et al. (2016)                    | PHC | IO          |                                               |
| 12     | Glenngård (2013)                                | PHC | IO          |                                               |
| 12     | McGarvey et al. (2019)                          | PHC | IO          | Data transformation                           |
| 13     | Rocha et al. (2014)                             | PHC | OO          | Input                                         |
| 14     | Ferrera et al. (2014)                           | PHC | IO          |                                               |
| 14     | Gouveia et al. (2016)                           | PHC | NO          |                                               |
Table 2 (continued)

| SubCAT | References                  | DMU  | Orientation | Undesirable outputs |
|--------|-----------------------------|------|-------------|---------------------|
| 14     | Razzaq et al. (2013)        | PHC  | OO          |                     |
| 15     | Yuan et al. (2020)          | PHC  | IO          |                     |
| 23     | Adang et al. (2016)         | PHC  | OO          | Data transformation |

*a COUN, country; HC, health care; NC/AC, nursing and aged care; PHC, primary health care; INS, insurance; LAB, laboratory; PAT, patient; PHY, physician
*b IO, input orientation; OO, output orientation; NO, no orientation; if blank, the corresponding information is not evident
*c WD, weak disposability; DEA, data envelopment analysis; NDEA, network DEA; DDF, directional distance function
*d The authors do not condition efficiency on quality variables; we therefore assigned their study to SubCAT 20.

Table 3 Classification of studies

| DMUs      | SubCAT | COUN | HR | HC | HC plus NC/AC | HC plus PHC | NC/AC | PHC | OTHER | Total |
|-----------|--------|------|----|----|---------------|-------------|-------|-----|-------|-------|
|           |        | 11   | 5  | 19 | 3             |             |       |     |       | 27    |
|           |        | 12   | 2  |    | 1             | 2           |       |     |       | 5     |
|           |        | 13   | 5  | 6  | 4             | 4           | 1     | 3   |       | 23    |
|           |        | 14   | 2  | 9  | 1             | 2           | 3     |     |       | 17    |
|           |        | 15   | 1  |    | 1             | 1           | 1     | 2   |       | 6     |
| Class 1   |        | 18   | 13 | 33 | 1             | 8           | 10    | 5   |       | 78    |
|           |        | 20   | 12 | 1  | 4             | 2           |       |     |       | 19    |
|           |        | 21   | 10 |    |               | 2           |       |     |       | 12    |
|           |        | 22   | 1  |    |               |             |       |     |       | 1     |
|           |        | 23   | 1  | 5  | 1             | 1           |       |     |       | 8     |
|           |        | 24   | 2  |    |               |             |       |     |       | 3     |
|           |        | 25   | 1  |    |               |             |       |     |       | 1     |
| Class 2   |        | 1    | 5  | 26 | 1             | 8           | 1     | 2   |       | 44    |
| CAT1      |        | 9    | 18 | 59 | 1             | 16          | 11    | 7   |       | 122   |
| CAT2      |        | 2    | 2  |    |               |             |       |     |       | 4     |
| Total     |        | 9    | 20 | 61 | 1             | 16          | 11    | 7   |       | 126   |

COUN, countries; HR, health regions; HC, hospital care; HC plus NC/AC, cross-provider assessment including hospital and nursing/aged care; HC plus PHC, cross-provider assessment including hospital and primary health care; NC/AC, nursing/aged care; PHC, primary health care; OTHER, all other providers
3.2 Studies by decision-making units

We built eight categories for DMUs (see Table 3). These include country studies (COUN); regional studies (health regions; HR); hospital studies (hospital care; HC); nursing and aged care studies (nursing care/aged care; NC/AC); studies devoted to the primary health care sector (primary health care; PHC); cross-provider studies covering the hospital and the nursing care sector (HC plus NC/AC) and the hospital and the primary health care sector (HC plus PHC); and a collective category (OTHER), which covers the remaining studies.

Hospital studies, including cross-provider studies, make up half of all studies (63 papers). Regional studies that investigate the efficiency of regions, provinces, districts, regional health boards and authorities (20 papers), nursing and aged care studies (16 papers), and primary health care studies (11 papers) together account for some 37% of all studies. Country studies (9 papers) and the residual DMU category “Other” (7 papers), which covers insurances, laboratories, physicians, and patients, make up the smallest DMU groups.

3.3 Studies over time

Figure 1 shows at least five studies per year since 2011. Before the year 2000, however, we identified only seven studies. These studies include those of Morey et al. (1992), Morey et al. (1995) and Ferrier and Valdmanis (1996), each investigating hospital efficiency; Kooreman (1994) and Kleinsorge and Karney (1992), who assess nursing efficiency; and Thanassoulis et al. (1995) and Salinas-Jiménez and Smith (1996), who evaluate health regions. There are no CAT2 studies before 2013.

Breaking down the evolution of studies by SubCATs (Fig. 2) shows an increase in

Fig. 1 Evolution of studies (Years with no studies are not displayed)
SubCATs 11, 13, and 14 since 2010. With two exceptions, NDEA papers (SubCATs 15 and 25) were published in 2019 and 2020. Before 2010 Class 2 articles are mainly SubCAT 20 studies. SubCATs 21 and 23 are more common in the second half of the last decade.

### 3.4 Studies by quality dimensions

Following Donabedian (1980), seven possible variants exist for taking quality into account in nonparametric efficiency studies. These include considering single dimensions (S, P, O) and the combinations thereof (SP, SO, PO, SPO). The figures in Table 4 read as follows. Only in Class 1 (78) and Stage 2 of Class 2 does the total number of quality variants correspond to the total number of papers for the following reasons. The 69 variants in Class 2 exceed the 44 Class 2 articles because these studies consider quality in both stages, except for SubCAT 20 articles. The number of variants in Stage 2 of Class 2 thus equals the 44 Class 2 papers. Two CAT2 studies condition the input–output space on quality indicators, in addition to including a priori a single quality indicator in the efficiency model (Gearhart and Michieka 2020) and investigating the association between a specific quality indicator and efficiency in the context of a second stage regression (Ferreira et al. 2013). Since we counted the quality dimensions used to condition the input–output space in both stages, the four CAT2 papers cover eight quality variants.

The last line in Table 4 reveals that outcome quality is by far the most frequently included single quality dimension (71). Considering combinations with outcome quality (i.e., SO, PO, SPO), the relevant figure increases to 112. Outcome quality appears as a single dimension in each Class 1 subcategory (46) and more than twice as often in Stage 1 of Class 2 studies (16) compared to Stage 2 (7).
Table 4 Quality dimensions by subcategory

| SubCAT | S  | P  | O  | SP | SO | PO | SPO | Total |
|--------|----|----|----|----|----|----|-----|-------|
| Class 1| 11 | 1  | 4  | 17 | 2  | 1  | 27  |       |
| 12     | 14 | 3  | 4  | 17 | 1  | 1  | 5   |       |
| 13     | 3  | 13 | 2  | 3  | 2  | 23 |     |       |
| 14     | 1  | 10 | 2  | 1  | 3  | 17 |     |       |
| 15     | 2  | 1  | 1  | 2  | 6  |    |     |       |
| Class 1| 1  | 8  | 46 | 7  | 1  | 6  | 78  |       |
| Class 2| 20 |    |    |    |    |    |     |       |
| Stage 1| 21 | 10 | 2  | 12 |    |    |     |       |
| 22     |    | 1  | 1  |    |    |    |     |       |
| 23     | 1  | 5  | 2  | 8  |    |    |     |       |
| 24     |    | 1  | 3  |    |    |    |     |       |
| 25     | 1  | 1  |    | 1  |    |    |     |       |
| Stage 1| 20 | 16 | 2  | 3  | 3  | 25 |     |       |
| Stage 2| 21 | 6  | 2  | 3  | 1  | 5  | 19  |       |
| 22     | 1  |    |    |    |    |    |     |       |
| 23     | 3  | 1  | 3  | 8  |    |    |     |       |
| 24     | 2  | 1  |    | 3  |    |    |     |       |
| 25     | 1  |    |    |    |    |    | 1   |       |
| Stage 2| 20 | 18 | 7  | 4  | 6  | 1  | 7   | 44    |
| Class 2| 18 | 2  | 23 | 6  | 9  | 4  | 7   | 69    |
| CAT1   | 19 | 10 | 69 | 13 | 10 | 13 | 13  | 147   |
| CAT2   | 1  | 2  | 1  | 2  | 2  | 8  |      |       |
| Total  | 20 | 10 | 71 | 13 | 11 | 15 | 15  | 155   |

S, structural quality; P, process quality; O, outcome quality; SP, structural and process quality; SO, structural and outcome quality; PO, process and outcome quality; SPO, structural, process and outcome quality

The studies frequently account for structural quality and combinations thereof in Stage 2. One explanation is that we found mixed handling of some structural quality indicators, such as teaching status. Some authors refer directly to quality when using these indicators, while others do not. To ensure a uniform approach, we classified them as structural quality indicators, regardless of whether the authors made direct reference to structural quality or not. Process quality is predominantly incorporated in Class 1 studies.

Breaking down the subcategories by DMU reveals the following (Table 5). In the country, regional, and hospital studies, outcome quality is the most frequently considered quality dimension. Structural quality is the predominant dimension in nursing and aged care studies, and process quality plays a significant role in primary health care studies. Hospital studies mainly account for outcome quality in the first stage and structural quality in the second stage. Nursing and aged care studies and primary health care studies incorporate outcome quality only in the efficiency model.
Table 5 Quality dimensions by subcategory and decision-making unit

| SubCAT | COUN | HR |
|--------|------|----|
|        | S    | P  | O  | SP | SO | PO | SPO | S | P  | O  | SP | SO | PO | SPO |
| Class 1 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 11   | 1  | 3  | 1  | 1  | 1  | 1   | 1 | 3  | 1  | 1  | 1  | 1  | 1   |
|         | 12   |    | 2  |    |    |    |     |   |    |    |    |    |    |     |
|         | 13   | 1  | 3  | 1  | 1  | 1  | 1   | 1 | 3  | 1  | 1  | 1  | 1  | 1   |
|         | 14   |    |    | 1  |    |    |     |   |    |    |    |    |    |     |
|         | 15   |    |    | 1  |    |    |     |   |    |    |    |    |    |     |
| Class 1 | 1    | 6  | 1  | 1  | 1  | 7  | 1   | 1 | 1  | 7  | 1  | 1  | 2  |     |
| Class 2 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
| Stage 1 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 20   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 21   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 22   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 23   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 24   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 25   |    |    |    |    |    |     |   |    |    |    |    |    |     |
| Stage 1 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 20   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 21   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 22   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 23   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 24   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 25   |    |    |    |    |    |     |   |    |    |    |    |    |     |
| Stage 2 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 20   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 21   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 22   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 23   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 24   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 25   |    |    |    |    |    |     |   |    |    |    |    |    |     |
| Stage 2 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 20   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 21   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 22   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 23   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 24   |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 25   |    |    |    |    |    |     |   |    |    |    |    |    |     |
| Class 2 |      |    |    |    |    |    |     |   |    |    |    |    |    |     |
|         | 2    |    |    |    |    |    |     |   |    |    |    |    |    |     |
Table 5 (continued)

| SubCAT | COUN | | | | | | HR | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|   | S | P | O | SP | SO | PO | SPO | S | P | O | SP | SO | PO | SPO |
| CAT1 | 1 | 8 | 1 | 2 | 3 | 12 | 1 | 1 | 2 | 2 |
| CAT2 | 1 | 2 | 1 | 1 | 2 | 1 |
| Total | 1 | 8 | 1 | 3 | 3 | 14 | 1 | 2 | 2 | 2 |

| SubCAT | HC* | | | | | | NC/AC | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|   | S | P | O | SP | SO | PO | SPO | S | P | O | SP | SO | PO | SPO |
| Class 1 | 11 | 3 | 14 | 1 | 2 |
| 12 | 4 | 1 | 2 | 1 |
| 13 | 1 | 6 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 1 |
| 15 | 4 | 24 | 1 | 4 | 1 | 1 | 4 | 1 | 2 |
| Class 1 Stage 1 | 20 | 9 | 1 | 1 | 1 |
| 21 | 22 | 1 | 1 | |
| 23 | 24 | 1 | 1 | 1 | 1 |
| 25 | Stage 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 |
### Table 5 (continued)

| SubCAT | HC* |   |   |   |   |   |   | NC/AC |   |   |   |   |   |   |   |
|------|-----|---|---|---|---|---|---|------|---|---|---|---|---|---|---|
|      | S   | P | O | SP | SO | PO | SPO | S   | P | O | SP | SO | PO | SPO |
| Stage 2 | 20  | 3 | 2 | 3 | 1 | 4 | 1 | 3   |   |   |   |   |   |   |
|         | 21  | 5 | 2 | 1 | 2 | 1 | 1 | 1   |   |   |   |   |   |   |
|         | 22  | 1 |   |   |   |   |   | 1   |   |   |   |   |   |   |
|         | 23  |   |   |   |   |   |   | 1   |   |   |   |   |   |   |
|         | 24  | 1 |   |   |   |   |   | 1   |   |   |   |   |   |   |
|         | 25  | 1 |   |   |   |   |   | 1   |   |   |   |   |   |   |
| Stage 2 | 11  | 4 | 5 | 1 | 6 | 4 | 4 | 4   |   |   |   |   |   |   |
| Class 2 | 11  | 14 | 7 | 2 | 6 | 4 | 2 | 5 | 1   |   |   |   |   |   |   |
| CAT 1  | 11  | 4 | 38 | 2 | 7 | 6 | 7 | 4 | 1 | 6 | 5 | 2 | 2 | 2   |
| CAT 2  | 2   | 2   |   |   |   |   |   | 2 | 2   |   |   |   |   |   |   |
| Total  | 11  | 4 | 38 | 2 | 7 | 8 | 9 | 4 | 1 | 6 | 5 | 2 | 2 | 2   |
| PHC    | S   | P | O | SP | SO | PO | SPO | S   | P | O | SP | SO | PO | SPO | Total |
| Class 1 | 11  | 1 | 2   |   |   |   |   | 27  |   |   |   |   |   |   |   |
|         | 12  | 1 | 1   |   |   |   |   | 5   |   |   |   |   |   |   |   |
|         | 13  | 1 |   | 1   |   |   |   | 1   | 1 | 1 | 1 | 1 | 1 | 1 | 23 |
|         | 14  | 2 | 1   |   |   |   |   | 17  |   |   |   |   |   |   |   |
|         | 15  | 1 |   |   | 1 | 1 | 1 | 6   |   |   |   |   |   |   |   |
| Class 1 | 1   | 4 | 4 | 1   | 1 | 1 | 2 | 1   | 78  |   |   |   |   |   |   |   |
| Class | Stage 1 | Stage 2 | Total |
|-------|---------|---------|-------|
|       |         |         |       |
| Class 1 | 1 | 4 | 5 |
| Class 2 | 6 | 1 | 7 |
| Class 3 | 8 | 1 | 9 |
| Class 4 | 10 | 1 | 11 |
| Class 5 | 12 | 1 | 13 |
| Class 6 | 14 | 1 | 15 |
| Class 7 | 16 | 1 | 17 |
| Class 8 | 18 | 1 | 19 |
| Class 9 | 20 | 1 | 21 |
| Class 10 | 22 | 1 | 23 |
| Class 11 | 24 | 1 | 25 |
| Class 12 | 26 | 1 | 27 |
| Class 13 | 28 | 1 | 29 |
| Class 14 | 30 | 1 | 31 |
| Class 15 | 32 | 1 | 33 |
| Class 16 | 34 | 1 | 35 |
| Class 17 | 36 | 1 | 37 |
| Class 18 | 38 | 1 | 39 |
| Class 19 | 40 | 1 | 41 |
| Class 20 | 42 | 1 | 43 |
| Class 21 | 44 | 1 | 45 |
| Class 22 | 46 | 1 | 47 |
| Class 23 | 48 | 1 | 49 |
| Class 24 | 50 | 1 | 51 |
| Class 25 | 52 | 1 | 53 |
| Class 26 | 54 | 1 | 55 |
| Class 27 | 56 | 1 | 57 |
| Class 28 | 58 | 1 | 59 |
| Class 29 | 60 | 1 | 61 |
| Class 30 | 62 | 1 | 63 |
| Class 31 | 64 | 1 | 65 |
| Class 32 | 66 | 1 | 67 |
| Class 33 | 68 | 1 | 69 |
| Class 34 | 70 | 1 | 71 |
| Class 35 | 72 | 1 | 73 |
| Class 36 | 74 | 1 | 75 |
| Class 37 | 76 | 1 | 77 |
| Class 38 | 78 | 1 | 79 |
| Class 39 | 80 | 1 | 81 |
| Class 40 | 82 | 1 | 83 |
| Class 41 | 84 | 1 | 85 |

Includes the cross-provider studies: HC plus NC and HC plus PHC; S, structural quality; P, process quality; O, outcome quality; SP, structural and process quality; SO, structural and outcome quality; PO, process and outcome quality; SPO, structural, process and outcome quality; COUN, countries; HR, health regions; HC, hospital care; PHC, primary health care; OTHER, all other providers.
3.5 Quality indicators by decision-making unit

This section highlights a few quality indicators, while the Appendix provides a complete list of the structural, process, and outcome indicators used in the reviewed articles. As the choice of quality indicators, among other things, heavily depends on the DMU category, we describe the indicators by DMU category. Given that outcome quality plays an essential role and is often directly included in the efficiency model, Table 2 also indicates the treatment of undesirable indicators in the efficiency model.

3.5.1 Country studies

Country studies belong to CAT1. Eight of nine country studies incorporate quality directly in the efficiency model (Class 1), while one study uses a two-stage approach (Class 2). Five of eight Class 1 studies are effectiveness studies (SubCAT 13), which means they only include quality outputs. These include, for example, life expectancy, disability-adjusted life expectancy, infant mortality, and potential years of life lost. One study additionally considers the share of infants immunized against measles as a process quality indicator (Ortega et al. 2017). Santos et al. (2012) investigate the process quality of health systems concerning preventing mother-to-child HIV infection. They use HIV tests for pregnant women up to appropriate therapies with antiretroviral medication for mothers and babies. Kim et al. (2020) use NDEA (SubCAT 15) to break down health system efficiency into health system investment and health system competitiveness. The authors incorporate the life expectancy, under-five mortality, and incidence of tuberculosis cases. Puertas et al. (2020) integrate healthy life years and infant mortality as outcome quality indicators in Stage 1 and investigate the variation in perceived health status using the effectiveness scores as explanatory variables in Stage 2 (SubCAT 23). Therefore, this study accounts for objective outcome quality in Stage 1 and investigates whether differences in subjective outcome quality are related to differences in effectiveness in Stage 2. Two studies assume that efficiency and quality represent different performance dimensions (Lo Storto and Goncharuk 2017; Kočišová et al. 2020) (SubCAT 12).

3.5.2 Regional studies

Effectiveness studies (SubCATs 13 and 23) are pretty widespread in the "HR" category. Class 1 (SubCAT 13) comprises assessments of Italian, Spanish, Mexican, and South African regions (Pelone et al. 2012; Carrillo and Jorge 2017; Serván-Mori et al. 2018; Ngobeni et al. 2020), UK health authorities (Salinas-Jiménez and Smith 1996), and Clinical Commissioning Groups (Takundwa et al. 2017). These studies differ considerably in terms of the considered quality dimensions.

Pelone et al. (2012) use the diabetes complication and coronary heart failure (CHF) hospitalization rates as disease-specific undesirable outcome indicators and the vaccination and referral rates as proxies for process quality. Serván-Mori et al. (2018) concentrate on process quality in providing maternal health services at the regional level, using indicators tailored to maternal care. Salinas-Jiménez and Smith (1996) and Takundwa et al. (2017) analyze the effectiveness of UK Health Boards.
and Clinical Commissioning Groups, respectively, albeit from different perspectives. Salinas-Jiménez and Smith (1996) investigate the structural and process quality of UK Health Boards, and Takundwa et al. (2017) concentrate on the outcome quality of Clinical Commissioning Groups. Typical for structural quality, Salinas-Jiménez and Smith (1996) incorporate characteristics involving General Practitioners (GPs) settings, such as the number of GPs per 10,000 listed patients, the share of GPs with less than 2,500 listed patients, and non-single-handed practicing GPs. They also use common process quality indicators, such as the share of GPs with higher targets for particular services, which also appeared in the country studies. Takundwa et al. (2017) incorporate objective outcome quality measures such as age- and sex-standardized survival rates, in addition to the willingness to recommend a GP practice as a subjective outcome quality indicator. Carrillo and Jorge (2017) also concentrate on objective and subjective outcome quality using healthy life expectancy at birth, infant survival, and self-perceived health status. Overall, the outcome indicators in the country and regional studies are pretty similar.

Five studies are Class 2 effectiveness studies (SubCAT 23), which differ in quality dimensions and indicators. Campanella et al. (2017), evaluating the efficiency of Italian hospital trusts, use disease-specific outcome quality indicators at both stages. While they include disease-specific mortality rates in their efficiency model, they use deficiencies in the outpatient care setting regarding avoidable hospital discharges and complication rates to explain efficiency differences among Italian hospital trusts. De et al. (2012) incorporate child immunization rates as process quality proxies in a second stage to examine the efficiency differences of Indian states. Amado and dos Santos (2009) and Allin et al. (2016) include objective outcome quality indicators in both stages and structural quality as an explanatory variable in Stage 2. Serván-Mori et al. (2020), in line with a previous study (Serván-Mori et al. 2018), investigate efficiency in providing maternal services at the regional level by including process quality indicators in the efficiency model. In Stage 2, they examine the variation in the maternal mortality ratio in Mexico, considering the efficiency from Stage 1 as an explanatory variable. Common to Class 2 effectiveness studies is that they account for structural quality indicators, if at all, in Stage 2.

SubCAT 11 is also a common subcategory. As regards their structure and process indicators, Giuffrida and Gravelle (2001) strongly rely on the quality indicators used by Salinas-Jiménez and Smith (1996). In addition, however, they also consider the number of patients registered with the GP as quantity output and the standardized mortality ratio as a nondiscretionary factor in the efficiency model to evaluate UK Family Health Authorities. Thanassoulis et al. (1995) were among the first to include both subjective (patient satisfaction) and objective outcome (babies at risk and their survival rates) quality indicators to assess the efficiency of UK District Health Authorities in providing perinatal care. Campoli et al. (2020) use common generic outcome indicators, such as life expectancy, child and maternal mortality, to investigate the efficiency of Brazilian Federative Units. Karsak and Karadayi (2017) use patient assessments of physical hospital characteristics (tangibility) and staff responsiveness (responsiveness) to derive two composite quality indexes. We categorized these proxies as outcome quality indicators (rather than structural and process quality indicators) because they represent patient views. Zheng et al. (2019) investigate the
efficiency of the Chinese health care system at the township and village levels. They use the average length of stay as a quality proxy at the township level. We assigned this indicator to structural quality because the authors related quality deficiencies to the lack of licensed doctors and advanced technologies in rural areas.

Du (2018), first developing a composite quality index, assesses the effect of quality on the efficiency of Chinese provinces by using models with and without the quality index (SubCAT 14). Since the composite index includes structural, process, and outcome quality elements, we assigned it to all three dimensions. In a similar approach, Polisena et al. (2010) investigate the efficiency of Canadian health integration networks using avoidable hospitalization rates.

Gearhart and Michieka (2020) condition the input–output space on quality-related attributes of the health care environment (personnel and facilities) (CAT2). They additionally consider the percentage of non-low weight births and the number of potential years of life gained below 75 years per 100,000 population as outcome quality indicators in the efficiency model. In evaluating the efficiency of Portuguese health regions, Ferreira et al. (2013) use a conditional order-m approach considering mortality rates. In addition, the authors investigate the association between efficiency and the share of patient complaints in the respective health region.

3.5.3 Hospital studies, including cross-provider studies

Hospital studies represent the largest DMU group, with no significant difference in the number of Class 1 and Class 2 articles. Studies in SubCat 11, the dominating subcategory in Class 1, do not incorporate structural quality. However, three studies include only process quality indicators. Miller et al. (2017) use disease-specific process indicators to investigate the effect of a US health care reform on hospital efficiency. Nedelea and Fannin (2017) explore efficiency differences between US hospitals against the background of their different reimbursement schemes. They use disease-specific clinical indicators similar to those of Miller et al. (2017). Matranga and Sapienza (2015) concentrate on organizational processes and use inappropriate discharges as proxies to evaluate the efficiency of Italian hospitals. Azadeh et al. (2016) first use a simulation approach to derive different scenarios for providing neurosurgical intensive care and then apply DEA to identify the most efficient scenario. The authors, using Iranian data, include mortality rates and bedsore patients as outcome quality indicators and average waiting days for operations and average postponed surgery plans as process indicators. Giménez et al. (2019) proxy process quality using the number of days waiting for an appointment and the minutes waiting for treatment in the emergency room. They use subjective (patient satisfaction) and objective indicators (readmissions and referrals) for outcome quality. The other SubCAT 11 studies concentrate on outcome quality.

5 In contrast to the full-frontier approaches, such as DEA, order-m and order-α are partial frontier methods. The input order-m frontier, for example, requires a subset of m peers to be drawn that produce at least as much of any output as the evaluated DMU. Among those peers, the one that uses the minimum input then serves as a benchmark to the DMU under evaluation. The computation of order-m efficiency involves solving an integral or resampling. In contrast, the order-α approach determines as a benchmark the input level that \((1 - \alpha) \times 100\text{ percent of the DMUs do not exceed to produce at least as much of any output as the DMU to be evaluated. The computation of order-α efficiency neither requires integral solving nor resampling. For details on partial frontier methods and computational aspects, see Daraio and Simar (2007a), Chapter 4.}
indicators, such as mortality (survival), readmission, infection, and referral rates. Park et al. (2018) use DEA twice. First, they derive a composite quality index using data from a patient survey covering satisfaction with medical care and service. Then they assess the efficiency of Korean hospitals using the quality index to adjust profits.

In a comparable study, Ferrier and Trivitt (2013) use DEA first to derive a process quality index, an outcome quality index, and an aggregate process and outcome quality index for selected diseases. Next, they examine separate models and use these indexes to adjust quantity outputs. In contrast to Park et al. (2018), they use various specifications so that their study falls into SubCAT 14. Navarro-Espigares and Torres (2011) apply an efficiency model without quality information, efficiency models with varying objective indicators entailing structural and process quality, and an effectiveness model with perceived quality (patient satisfaction) as a single output. Nayar and Ozcan (2008) incorporate only process quality indicators to analyze the efficiency of US hospitals. The remaining studies in SubCAT 14 include common outcome quality indicators.

With four studies, hospital effectiveness studies (SubCAT 13) are less common. These studies incorporate outcome quality indicators, with the majority relying on objective indicators, such as dead (live) hospital discharges, disease-specific readmission rates, and mortality and infection rates. However, Kholghabad et al. (2019) use patient satisfaction to investigate the effectiveness of Iranian emergency departments. In a single SubCAT 15 study, Cohen-Kadosh and Sinuany-Stern (2020) account for process and outcome quality indicators tailored to assessing the surgical treatment of hip fracture patients. They use NDEA to investigate the efficiency of orthopedic departments in acute care Israeli hospitals.

The SubCATs 20 and 21 studies are relatively balanced. In SubCAT 20, by definition, quality is not considered in Stage 1. With one exception (Dutta et al. 2014), the studies in SubCATs 21 to 25 account for structural quality only in Stage 2. Frequently used structural quality indicators include teaching status, certification and accreditation, technology status, and the availability of qualified personnel (e.g., registered nurses). Some studies tailored the structural quality indicators to the respective research question. For example, Huerta et al. (2011) rely on computerized physician order entries to examine the association between health information technology and US hospital efficiency. Process indicators appear in Stage 2, mainly because we assumed that the composite indexes such as certification and accreditation indexes contain elements of process quality. The outcome quality indicators used in SubCATs 20 and 21 are comparable to those in Class 1. Ferreira et al. (2020), who incorporate composite indexes capturing appropriateness, timeliness, and accessibility of care in a second stage, justify their two-stage approach based on their previous findings that the separability condition holds for Portuguese hospitals (see Ferreira and Marques 2019).

Roth et al. (2019), using a two-stage approach, are the only ones who assume that quality and efficiency represent different dimensions of hospital performance.

---

6 We assumed that indicators such as certification and accreditation include elements of all quality dimensions.

7 The authors also use a conditional approach. In line with the categorization of Class 2 studies (see Sect. 2.2), we assigned this article to CAT1 as the authors do not condition efficiency on quality.
**SubCAT 22.** SubCAT 24 and SubCAT 25 are rare. Chiu et al. (2008) incorporate structural quality at both stages, using medical equipment and staff as proxies. They also consider the doctor’s time spent for diagnosis, categorized as a process quality indicator. Li (2015) examines the relationship between mortality and efficiency using mortality measures from different aggregation levels. Khushalani and Ozcan (2017)\(^8\) use a dynamic NDEA to analyze the evolution of efficiency and quality of US hospitals. They decompose the care process into a surgical/medical subprocess (including only quantity outputs\(^9\)) and a quality subprocess (including subjective and objective disease-specific outcome quality indicators).

Two more studies show a connection to the hospital sector (HC plus NC/AC and HC plus PHC). Laine et al. (2005) investigate efficiency in providing long-term nursing care services in hospitals and nursing care facilities. The authors incorporate indicators covering all quality dimensions and consider them to affect the distribution of inefficiency scores (SubCAT 20). Their process and outcome quality indicators differ considerably from those used in hospital efficiency studies, as their focus is on maintaining and restoring the physical and mental health of the elderly rather than providing acute care. However, the authors offer a rich insight into quality indicators suitable for incorporation into nursing care studies. Obure et al. (2016) connect hospital care with that of primary care. The study relies on models with and without quality indicators (SubCAT 14). It includes composite quality indexes that capture structural and process quality elements in assessing the efficiency of various health facilities (e.g., hospitals, clinics, and other health facilities, centers, and units offering services for HIV patients and their sexual and reproductive health).

Two CAT2 studies relate to the hospital sector. Varabyova et al. (2017) use a conditional order-m approach to assess the efficiency of German cardiology departments. They investigate the impact of process quality (patient radiation exposure) and subjective (patient satisfaction), and objective (mortality ratio) outcome quality indicators on the input–output space. Ferreira and Marques (2019) choose a conditional order-\(\alpha\)\(^10\) approach to analyze the effect of quality indicators covering elements of all quality dimensions on the efficiency of Portuguese hospitals.

### 3.5.4 Nursing and aged care studies

Researchers started to look at the quality of nursing care quite early. The first studies date back to the early 1990s. The number of Class 1 and Class 2 studies is balanced. As for effectiveness studies (SubCAT 13), Bott et al. (2007) and Shen et al. (2020) concentrate on organizational processes of US nursing facilities and Korean

---

8 However, Khushalani and Ozcan (2017) could also be a SubCAT 15 study because the authors themselves do not explicitly declare the indicator “teaching status” as a structural quality proxy. Rather, we recorded the teaching status as a structural quality indicator for the sake of consistency. See also Sect. 3.4 for a discussion of this issue.

9 The authors also incorporate a registered-to-total-nurses ratio and a high-technology index as input for evaluating the surgical/medical subprocess. As similar indicators are used by other authors as proxies for structural quality, we also classified this study as one that covers the dimension of structural quality.

10 See FN 5.
halfway houses, respectively. Shwartz et al. (2016) and Min et al. (2016) address outcome quality and use care-specific indicators, such as patients with pressure ulcers and other mental and physical restraints. Shwartz et al. (2016) additionally consider risk-adjusted mortality as an outcome quality indicator to analyze the efficiency of US nursing homes. Shimshak and Lenard (2007) use separate models to assess US nursing homes, accounting for outcome quality considering residents with indwelling or external catheters, pressure sores or ulcers, and other restraints (SubCAT 12). Another two studies cover structural quality. Avkiran and McCrystal (2014) incorporate the average length of service of registered nurses in their NDEA approach (SubCAT 15), thereby capturing the quality-related characteristics of the nursing care setting. Kleinsorge and Karney (1992) use a state inspection score and the share of FTEs per resident as indicators of the quality of the care setting (SubCAT 14).

In Class 2, a heterogeneous picture emerges in the corresponding subcategories. Nursing studies consider structural and process quality predominantly in Stage 2 and outcome quality exclusively in Stage 1. Lee et al. (2009), Lin et al. (2017), and Ni Luasa et al. (2018) investigate the association between efficiency and structural and process quality using various ratings of the nursing care setting and the care processes (SubCAT 20). However, Zhang et al. (2008) include quality-adjusted outputs, which contain structural and process quality elements, in their efficiency model (SubCAT 24), and Garavaglia et al. (2011) use extra nursing hours as proxies for structural quality in Stage 1 (SubCAT 21). No Class 2 study accounts for outcome quality in Stage 2. Wichmann et al. (2018) tackle the quality of dying in their effectiveness study, with tailored outcome quality indicators in Stage 1 and structural quality indicators in Stage 2 (SubCAT 23).

3.5.5 Primary health care studies

Similar to country studies, Class 1 studies dominate in the primary health care category. In SubCAT 11, Aduda et al. (2015) look at the efficiency of providing voluntary medical circumcision for men in Kenyan health facilities using a composite quality index in their efficiency model. Milliken et al. (2011) adjust quantity outputs by using various composite quality indicators in evaluating the efficiency of Canadian primary care practices. Ruiz-Rodriguez et al. (2016) account for process quality in assessing the efficiency of providing services related to women’s health prevention and promotion programs in Columbia. Rocha et al. (2014) evaluate management practices in Brazilian family health teams using indicators related to organizational processes (SubCAT 13). In SubCAT 14, Ferrera et al. (2014) rely on a composite quality index with structural and process quality elements to assess the efficiency of primary care centers in Spain. Gouveia et al. (2016) and Razzaz et al. (2013) address patient satisfaction to identify best practices in Portuguese and Iranian primary care centers, respectively.

In contrast to these studies, Glennsgard (2013) and McGarvey et al. (2019) treat quality and efficiency as separate performance dimensions (SubCAT 12). The former assesses Swedish primary care centers. The latter focuses on prenatal care provided in US health centers, considering tailored process (percentage of prenatal patients receiving care in the first trimester) and outcome (percentage of low birth weight births) quality indicators. Yuan et al. (2020) rely on composite quality indexes to evaluate
Chinese community health centers as proxies for outcome quality within a NDEA framework (SubCAT 15). In a single Class 2 study, Adang et al. (2016) examine the effectiveness of Dutch primary health care in preventing cardiovascular diseases by focusing on cardiovascular risk management (SubCAT 23). They incorporate disease-specific process and outcome quality indicators in Stage 1 and proxies for structural quality in Stage 2.

3.5.6 Other studies

This collective DMU category comprises physician effectiveness studies (Dowd et al. 2014; Fiallos et al. 2017), insurance efficiency studies (Bryce et al. 2000; Villamil et al. 2017), and medical laboratory efficiency assessments (Ghafari Someh et al. 2020; Lamovšek and Klun 2020). A single article evaluates the quality of life of Diabetes Type 1 patients (Borg et al. 2019). Due to their heterogeneity, we refrain from presenting the studies’ quality indicators but refer the reader to the Appendix for further information.

4 Discussion

The decision to account for specific quality dimensions and indicators in efficiency studies results from weighing potential advantages and disadvantages. We, therefore, assume that the differences in the quality dimensions and indicators explain at least in part the breadth of the methodological approaches.

Quentin et al. (2019, p. 40) discuss the pros and cons of the Donabedian quality framework, including the corresponding indicators. They note that structural quality is "easily available, reportable and verifiable". The frequent use of indicators such as teaching, accreditation and certification status, and specific technology and staff ratios in hospital studies supports this statement. Nursing care studies also use equivalent proxies. For example, Kleinsorge and Karney (1992) incorporate a state inspection score in their assessment.

Most studies that incorporate structural quality as a single quality dimension in Stage 2 (Class 2) underpin the assessment that "the link between structures and clinical processes or outcomes is often indirect and dependent on the actions of health care providers" (Quentin et al. 2019, p. 40). Two CAT2 studies investigate this issue. Ferreira and Marques (2019) analyze the impact of access and quality indicators on efficiency, while Gearhart and Michieka (2020) concentrate on some staff categories.

However, the frequent use of structural quality indicators in Stage 2 may result from classifying several indicators as proxies for structural quality that some authors did not explicitly associate with this quality dimension. This finding can also indicate the rather diffuse relationship between structural quality on the one hand and process and outcome quality or efficiency on the other (Quentin et al. 2019, p. 38).

Process indicators are objective indicators that are also easy to survey and interpret. In addition, there is a direct link between process and outcome quality. NDEA models are advanced approaches to investigate this link. Cohen-Kadosh and Sinuany-Stern (2020) incorporate the share of patients with surgery within 48 h and drain usage for
one day as process indicators in their NDEA model to evaluate the surgical treatment of hip fracture patients in Israeli orthopedic hospital departments.

The main advantage of process indicators is that compliance with process indicators is easy to check, and the assignment of misconduct to the responsible person is straightforward. Process indicators are used to study organizational and medical aspects of service delivery. Disease-specific measures prevail in the latter. We identified both types in the reviewed papers. Rocha et al. (2014) use the share of scheduled queries among team queries and the percentage of actions developed to investigate management practices in Brazilian family health teams. Amado and Dyson (2009) incorporate patients with a diabetes annual review in the previous 12 months, patients receiving antihypertensive medication, and patients under regular treatment with lipid-lowering medicines to assess the efficiency of primary care practices providing diabetes care.

However, process indicators also have drawbacks. As the name suggests, process indicators relate to single processes. Studies assessing efficiency at a highly aggregated level, such as at the hospital level, cannot include all available process indicators. In this regard, we identified one of the following approaches: a focus on selected services or service providers, a focus on a few selected disease-specific indicators, the use of a composite process index, or a combination thereof. With the percentage of prenatal patients who receive care in the first trimester, McGarvey et al. (2019) concentrate on the prenatal services provided at US federally-qualified health centers. Miller et al. (2017) assess the effect of hospital reform on total hospital costs and quality by focusing on three disease-specific process indicators (heart attack, heart failure, and pneumonia). Ferrera et al. (2014) incorporate a composite quality index to assess the efficiency of Spanish primary care centers. Obure et al. (2016) also rely on a quality index to examine the effectiveness of HIV services and services strengthening patients’ sexual and reproductive health.

Additionally, a dimensionality problem may arise with the use of process indicators. Nonparametric efficiency studies in particular often evaluate comparatively few DMUs. However, with many process indicators and few DMUs, it is challenging to identify efficiency differences. Using a few condition-specific process indicators and composite quality indexes are possible alternatives. Neither of the two variants is free from disadvantages. While the concentration on some conditions reflects only a small part of process quality, composite indexes may lead to a loss of information.

Outcome quality and its visibility are becoming increasingly important. Outcome quality is particularly relevant to the patients, as it reflects treatment success and affects their wellbeing. The outcome indicators used in the studies are thoroughly homogeneous within the corresponding DMU category. However, the use of outcome quality indicators entails several particularities.

First, we observed a heterogeneous picture for studies that directly included undesirable outcome indicators in the efficiency model. While some authors treat undesirable outcome indicators as inputs (e.g., Guo et al. 2017), others apply a linear transformation (e.g., Adang et al. 2016), while still others transform them a nonlinear way (e.g., Wang et al. 2018). Assuming weak disposability for undesirable outcome indicators is another possibility (e.g., Prior 2006). Still, others use sophisticated metrics such as directional distance functions or hyperbolic measures, which do not require a transformation of undesirable outputs (e.g., Giménez et al. 2019; Villamil et al. 2017).
Second, the studies use objective or subjective indicators as proxies of outcome quality. However, objective indicators are usually not patient invariant and require risk adjustment to capture heterogeneities in the patient population. Using risk-adjusted mortality for three different conditions, Campanella et al. (2017) do just that in assessing the efficiency of Italian hospital trusts. Subjective outcome measures, such as patient satisfaction, are also increasingly being incorporated in efficiency studies. Relevant studies, among other things, examine the relationship between subjective outcome indicators and efficiency (e.g., Barpanda and Sreekumar 2020).

Outcome quality indicators are by far the most widely used indicators, despite the challenges of correctly establishing a direct and immediate causal relationship between a service provider’s actions and the effect on the patient’s condition (Quentin et al. 2019, p. 53). In hospital efficiency studies, for example, mortality rates are pretty standard outcome quality indicators. However, several factors beyond the service provider’s control can affect hospital mortality, including pre-hospital care, patient characteristics, and other environmental factors. Delays in the treatment effects on the patient’s health are also an issue. Studies already account for the patient characteristics, delays in treatment effects, and heterogeneous environments. For example, Ferrier and Trivitt (2013) include condition-specific post-discharge mortality rates in assessing US hospitals, and Ferreira et al. (2020) account for various environmental factors in evaluating the efficiency of Portuguese hospitals.

5 Summary and limitations

The review shows a variety of methods to integrate quality into efficiency studies. One reason for the diverse approaches is certainly the 40-year observation period. Over the years, scholars continued to introduce sophisticated techniques into the literature. Apart from the considerable variation in DMU-types and their aggregation level, the model heterogeneity also results from the differences in research questions. In contrast to Kohl et al. (2018), who assign DEA hospital studies to four broad research questions, the present review does not cluster by DMU-specific research questions. Further clustering would not have produced meaningful results due to comparatively few studies in some DMU categories combined with quite heterogeneous research questions. The model diversity is undoubtedly also due to country-specific features. The reviewed studies cover a broad spectrum of different health systems—from emerging countries with underdeveloped (public) health systems to highly developed countries with comprehensive public health care, where well-trained medical professionals use advanced medical technology. Accordingly, various specifications emerged. Finally, differences in the availability and accessibility of relevant quality data can also explain variations in model specifications.

Regarding the quality dimensions, the review shows that many studies assume that structural quality affects inefficiency distribution. At the same time, they presume that process and outcome quality predominantly impact the efficiency frontier. These findings, however, are subject to several limitations.

First, the choice of the database and keywords may affect the inclusion of studies in the review. The review does not cover papers published in journals not included
in the Science Citation Index Expanded, the Social Sciences Citation Index, or the Emerging Sources Index. It does not cover articles with keywords other than those used for the review either. Second, given the recently introduced conditional approaches, we expect a shift in publications favoring CAT2 in the following years due to the dynamic publication activity. Third, the assignment of quality indicators to quality dimensions entails a considerable degree of subjectivity that also carries over to the conclusions drawn in this review. Consequently, we disclose the assignment of the quality indicators to the quality dimensions in the Appendix.

Overall, the review does not recommend how to proceed if one intends to incorporate quality into efficiency analysis. The review’s findings highlight that this would be challenging even within a single DMU type. The decision for a model specification depends on many factors, such as the specific research question, the environmental conditions, and the data availability. Instead, the overview of past efficiency approaches combined with the extensive collection of quality indicators by DMU-specific quality dimension can serve as a starting point for future efficiency studies aimed at integrating quality.

Funding  Open access funding provided by University of Graz. No funds, grants, or other support was received.

Open Access  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/.

Appendix

See Table 6.
### Table 6 Quality indicators by quality dimension and subcategory

| SubCAT | References | DMU | Stage 1 | Outcome |
|--------|------------|-----|---------|---------|
| 12     | Kočišová et al. (2020) | COUN | Structural | Standardized survival rates of CVD patients |
| 12     | Lo Storto and Goncharuk (2017) | COUN | Process | Infant mortality per population; HLYs at birth; life expectancy at birth |
| 13     | Adang and Borm (2007) | COUN | Outcome | Life expectancy at birth; infant mortality |
| 13     | Althin et al. (2019) | COUN | Structural | Age-standardized 5-year relative survival rates (lung cancer, breast cancer); disability-adjusted life expectancy |
| 13     | Kozuñ-Ciešlak (2020) | COUN | Process | Infant mortality per 1,000 live births; PYLL per 1,000 inhabitants |
| 13     | Ortega et al. (2017) | COUN | Outcome | % 1-year-old children with immunization against measles |
| 13     | Santos et al. (2012) | COUN | Structural | Under-five mortality per 1,000 live births |

Pregnant women tested for HIV; pregnant women living with HIV, who received antiretrovirals for preventing mother-to-child transmission; infants born to women living with HIV receiving antiretrovirals; infants born to women living with HIV receiving cotrimoxazole prophylaxis within two months of birth.
| SubCAT | References | DMU  | Stage 1 | Outcome |
|--------|------------|------|---------|---------|
|        |            |      | Structural | Process | Outcome |
| 15     | Kim et al. (2020) | COUN | Life expectancy at birth; incidence of tuberculosis per 100,000 population; under-five mortality rates per 1,000 live births |
| 23     | Puertas et al. (2020) | COUN | HLY’s at birth; infant mortality per 1,000 live births |
| 11     | Arrieta and Guillén (2017) | HC | Risk-adjusted perinatal mortality rates; risk-adjusted perinatal near miss |
| 11     | Azadeh et al. (2016) | HC | Avg waiting days for operation (post-admission); avg postponed surgery plans (post-admission) | Avg mortalities; avg bedsore patients |
| 11     | Bilsel and Davutyan (2014) | HC | Deaths per surgery (alternatively, deaths per inpatient) |
| 11     | Chang et al. (2011) | HC | Net inpatient mortality (mortalities occurring 48 h after ambulatory procedures without ER mortalities, stillbirths, and early neonatal mortalities) |
| 11     | Chen et al. (2016) | HC | Net inpatient mortalities; referrals; nosocomial infections |
| SubCAT | References                  | DMU   | Stage 1                                                                 | Outcome                                                                 |
|--------|-----------------------------|-------|-------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 11     | Chen et al. (2019)          | HC    |                                                           | Net inpatient mortality; referrals; nosocomial infections               |
| 11     | Du et al. (2014)            | HC    |                                                           | Weighted avg in-hospital survival rates for 16 common conditions        |
| 11     | Giménez et al. (2019)       | HC    | Days waiting for an appointment; minutes waiting in ER                  | Satisfied patients; readmissions within 20 days; patients referred to higher levels |
| 11     | Guo et al. (2017)           | HC    |                                                           | Crude mortality rates                                                  |
| 11     | Hu et al. (2012)            | HC    |                                                           | Risk-adjusted patient mortality                                         |
| 11     | Kalis and Stracova (2019)   | HC    |                                                           | Mortality                                                              |
| 11     | Matranga and Sapienza (2015)| HC    | Inappropriate discharges; inappropriate day-hospital discharges         |                                                                         |
| SubCAT | References              | DMU | Structural | Process                                                                 | Outcome                                                                 |
|--------|-------------------------|-----|------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 11     | Miller et al. (2017)    | HC  |            | % HA patients: given ACE inhibitor or ARB for left ventricular systolic dysfunction, given aspirin upon arrival; prescribed aspirin at discharge, prescribed beta-blocker at discharge, given smoking cessation counseling; % HF patients: given an evaluation of left ventricular systolic function, given ACE inhibitor, given smoking cessation advice/counseling, discharge instructions; % PN patients: given antibiotics within 6 h of arrival, administered ER blood culture before given antibiotics, given smoking cessation counseling, given pneumococcal vaccination, given most appropriate initial antibiotics | 30-day risk-adjusted actual mortality; 30-day risk-adjusted forecast mortality |
| 11     | Morey et al. (1995)     | HC  |            |                                                                         |                                                                         |
| 11     | Nedelea and Fannin (2017)| HC  |            | % PN patients with appropriate initial antibiotic treatment; % HF patients with an assessment of the LVS function |                                                                         |
| SubCAT | References | DMU | Structural | Process | Outcome |
|--------|------------|-----|------------|---------|---------|
| 11     | Park et al. (2018) | HC  |            |         | Quality-adjusted profit (medical service; service level satisfaction) |
| 11     | Sinuany-Stern et al. (2016) | HC  |            |         | Deaths |
| 11     | Thanassoulis et al. (2016) | HC  |            |         | Time to patient readmission from the discharge of the assessed episode |
| 11     | Wu et al. (2013) | HC  |            |         | CMI-adjusted readmissions |
| 13     | Chang and Lan (2010) | HC  |            |         | Mortality rates within one month; infection rates within one month |
| 13     | Dismuke and Sena (2001) | HC  |            |         | Dead discharges; live discharges |
| 13     | Kholghabad et al. (2019) | HC  |            |         | Satisfaction with: treatment time, emergency environment, doctors’ performance, nurses’ performance, equipment quality |
| 13     | Williams et al. (2016) | HC  |            |         | HF and PN: 30-day hospital readmission rates, deaths within 30 days from admission |
| Stage 1 | Structural | Process | Outcome |
|---------|------------|---------|---------|
| SubCAT | References | DMU     |         |
| 14      | Ferrier and Trivitt (2013) | HC      | Process quality index (AMI patients: receiving aspirin within 24 h of arrival, prescribed aspirin at discharge, prescribed beta-blocker at discharge, receiving beta-blocker within 24 h of arrival; CHF patients: prescribed ACE inhibitor at discharge, with documented LVF assessment or one scheduled at discharge; PN patients: with initial blood culture collected before given antibiotics, receiving first antibiotics within 24 h of arrival) Aggregate process and outcome quality index | Outcome quality index (risk-adjusted mortality rates within 30 days of discharge for AMI, within 180 days for CHF, within 30 days for PN; risk-adjusted surgical complication rates for cholecystectomy) Aggregate process and outcome quality index |
| 14      | Karagiannis and Velentzas (2012) | HC      | % survival after admission |
| 14      | Navarro-Espigares and Torres (2011) | HC      | CMI (∑ CMI-adjusted DRGs per discharge); % programmed surgical interventions; surgical turnover (hours surgical rooms are used per total hours available); diagnoses at discharge; procedures at discharge; ALOS; expected ALOS; pre-operative ALOS; stays utilization index; % Caesarean births | CMI (∑ CMI-adjusted DRGs per discharge); % programmed surgical interventions; surgical turnover (hours surgical rooms are used per total hours available); diagnoses at discharge; procedures at discharge; ALOS; expected ALOS; pre-operative ALOS; stays utilization index; % Caesarean births | Patient satisfaction |
| SubCAT | References                                  | DMU | Structural                                                                 | Process                                                                                   | Outcome                                                                                           |
|--------|---------------------------------------------|-----|---------------------------------------------------------------------------|------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| 14     | Nayar and Ozcan (2008)                      | HC  | % PN patients: receiving first antibiotic treatment within 4 h of arrival, given oxygenation assessment, 65 + given (advice for) pneumococcal vaccination | Level of nosocomial infections                                                              | Mortality rates                                                                                   |
| 14     | Prior (2006)                                 | HC  |                                                                            | Clinically active infections per patient                                                   |                                                                                                   |
| 14     | Sola and Prior (2001)                       | HC  |                                                                            | Mortality rates                                                                            |                                                                                                   |
| 14     | Tiemann and Schreyögg (2009)                 | HC  |                                                                            |                                                                                           |                                                                                                   |
| 14     | van Ineveld et al. (2016)                   | HC  |                                                                            | Patients with decubitus ulcers                                                              |                                                                                                   |
| 14     | Yang and Zeng (2014)                        | HC  |                                                                            | Mortality rates; ALOS rates                                                                 |                                                                                                   |
| 15     | Cohen-Kadosh and Sinuany-Stern (2020)       | HC  | % patients with surgery within 48 h, using a drain for one day             | % patients: who died within 365 days of surgery, with wound infections within 365 days of surgery |                                                                                                   |
| 20     | Chu et al. (2002)                           | HC  |                                                                            |                                                                                           |                                                                                                   |
| 20     | Ferreira et al. (2020)                      | HC  |                                                                            |                                                                                           |                                                                                                   |
| 20     | Ferrier and Valdmanis (1996)                | HC  |                                                                            |                                                                                           |                                                                                                   |
| 20     | Gautam et al. (2013)                        | HC  |                                                                            |                                                                                           |                                                                                                   |
| 20     | Gholami et al. (2015)                       | HC  |                                                                            |                                                                                           |                                                                                                   |
| 20     | Gok and Sezen (2013)                        | HC  |                                                                            |                                                                                           |                                                                                                   |
| SubCAT | References | DMU | Stage 1 |
|--------|------------|-----|---------|
|        |            |     | Structural | Process | Outcome |
| 20     | Harrison and Coppola (2007) | HC |          |          |         |
| 20     | Huerta et al. (2011) | HC |          |          |         |
| 20     | Kittelsen et al. (2015) | HC |          |          |         |
| 20     | Lindlbauer et al. (2016) | HC |          |          |         |
| 20     | Mujasi et al. (2016) | HC |          |          |         |
| 20     | Pai et al. (2019) | HC |          |          |         |
| 21     | Barpanda and Sreekumar (2020) | HC |          |          | Mortality rates |
| 21     | Clement et al. (2008) | HC |          |          |         |
| 21     | Dutta et al. (2014) | HC |          | Quality-adjusted (using a doctor involvement index) | outpatients treated per day |
| 21     | Leleu et al. (2018) | HC | |          | 30-day risk-adjusted mortality: HF, HA, PN; 30-day risk-adjusted readmission rates: HF, HA, PN, hip and knee, all other |
| 21     | Lin et al. (2019) | HC | |          | Net inpatient survivors |
| SubCA | References | DMU | Structural | Process | Outcome |
|-------|------------|-----|------------|---------|---------|
| 21    | Morey et al. (1992) | HC  |            |         | Forecasted mortalities; actual mortalities |
| 21    | Mutter et al. (2010) | HC  |            |         | Failure to rescue; infection due to medical care; postoperative respiratory failure; postoperative sepsis |
| 21    | Nayar et al. (2013) | HC  |            |         | Survival rates: AMI, CHF, PN |
| 21    | Valdmanis et al. (2008) | HC  |            |         | Failure to rescue; infection due to medical care; postoperative respiratory failure; postoperative sepsis |
| 21    | Wang et al. (2018) | HC  |            |         | In-hospital standardized mortality ratios |
| 22    | Roth et al. (2019) | HC  | Process of care index |         | % survival; % not-readmitted; satisfaction with: communication, cleanliness, and quietness; responsiveness to patient needs; overall satisfaction; recommend hospital |
| 24    | Chiu et al. (2008) | HC  | Patients serviced by expensive medical equipment | Doctor's avg number of outpatients | |
| 24    | Li (2015) | HC  |            |         | Risk-adjusted mortality ratio: DRG level; level of major diagnoses |
| SubCAT | References | DMU | Stage 1 | Outcome |
|--------|------------|-----|---------|---------|
|        |            |     | Structural | Process |         |
| 25     | Khushalani and Ozcan (2017) | HC | RNs ratio; hi-tech service mix | 30-day risk-standardized mortality rates: HA, HF, PN; 30-day risk-standardized readmission rates: HA, HF, PN; % patients who would definitely recommend hospital; % patients who gave hospital high rating |
| 30     | Ferreira and Marques (2019) | HC | (Hospital beds per 1,000 inhabitants; FTE nurses per 1,000 standard patients; FTE doctors per 1,000 standard patients; inpatient bed occupancy rates; OR capacity utilization)* | (Outpatient surgeries on potential outpatient procedures; waiting time before surgery; rates of first medical appointment within time; rates of surgeries within time; hip fracture surgery in first 48 h)* | (Readmission rates within 30 days after discharge; rates of inpatients staying more than 30 days; decubitus ulcer rates; postoperative pulmonary embolism/deep vein thrombosis rates; postoperative septicemia rates; obstetric trauma rates (vaginal delivery with an instrument); obstetric trauma rates (vaginal delivery without instrument); in-hospital death rates for low severity cases; stillbirth rate; childhood mortality rates)* |
| 30     | Varabyova et al. (2017) | HC | (Patient radiation exposure)* | (Patient satisfaction; standardized mortality ratio)* |
| SubCAT | References | DMU | Stage 1 | Process | Outcome |
|--------|------------|-----|---------|---------|---------|
| 20     | Laine et al. (2005) | HC plus NC/AC | Structural | Process | Outcome |
| 14     | Obure et al. (2016) | HC plus PHC | Composite quality index; structural quality index | Composite quality index; process quality index | |
| 11     | Campoli et al. (2020) | HR | | | Child mortality per 1,000 live births; maternal mortality (deaths up to 42 days after the termination of pregnancy); life expectancy |
| 11     | Giuffrida and Gravelle (2001) | HR | GPs; practices with nurses; GPs < 2,500 listed patients; GPs not practicing single-handed; practices satisfying minimum standard of Statement of Fees and Allowances; practice nurses; GPs on the minor surgery list | GPs achieving 80% target for cervical cytology testing; GPs achieving a higher target for childhood immunization; GPs achieving a higher target for pre-school boosters | Mortality ratio |
| 11     | Karsak and Karadayi (2017) | HR | | | Tangibility; responsiveness |
| 11     | Thanassoulis et al. (1995) | HR | | | Very satisfied mothers; satisfied mothers; babies at risk; survival rates of babies at risk |
| 11     | Zheng et al. (2019) | HR | ALOS in township hospital | | |
### Table 6 (continued)

| SubCAT | References                          | DMU   | Stage 1                                                                 | Outcome                                                                                     |
|--------|------------------------------------|-------|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| 13     | Carrillo and Jorge (2017)          | HR    | Health care quality in nonparametric efficiency studies: a...           | Healthy life expectancy at birth (years); infant survival rates; % population with positive self-perceived health status |
| 13     | Ngobeni et al. (2020)              | HR    |                                                                         | Infant mortality rates                                                                      |
| 13     | Pelone et al. (2012)               | HR    |                                                                         | Influenza vaccination rates (per 100 population 65 +); referral rates to secondary care (per 100,000 population) |
| 13     | Salinas-Jiménez and Smith (1996)   | HR    | GPs per 10,000 listed patients; % practices with practice nurse; % GPs with less than 2,500 listed patients; % non-single-handed practicing GPs; % premises satisfying minimum requirements set out in the Statement of Fees and Allowances | Long term complication diabetes hospitalization rates (per 1,000 population 18 +); CHF hospitalization rates (per 1,000 population 65 +) |
| 13     | Serván-Mori et al. (2018)          | HR    |                                                                         | Antenatal care during pregnancy; timely antenatal care; frequent antenatal care; institutional care during childbirth; medical care during childbirth; adequate maternal care |
| SubCAT | References          | DMU  | Stage 1                                                                 |
|--------|---------------------|------|--------------------------------------------------------------------------|
|        |                     |      | Structural Process Outcome                                               |
| 13     | Takundwa et al. (2017) | HR   | Standardized avg health status scores 18 + ; Age and sex standardized survival rates: respiratory diseases, cancer; Achievement score: cardiovascular diseases, cancer, COPD; % patients who would recommend GP practice to others |
| 14     | Du (2018)           | HR   | Composite quality index                                                 |
| 14     | Polisena et al. (2010) | HR   | Composite quality index                                                 |
| 23     | Allin et al. (2016)  | HR   | Hospitalization rates (age-standardized inpatient acute care for conditions where proper ambulatory care prevents the need for hospitalization) |
|        |                     |      | Survival rates from treatable causes (< age 80); PYLL from treatable causes (< age 80); PYLL from treatable causes (< age 85); PYLL from treatable causes (< age 75); mortality from treatable causes (< age 80) |
| SubCAT | References                          | DMU  | Stage 1                                                                 | Outcome                                                                 |
|--------|------------------------------------|------|-------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 23     | Amado and Dyson (2009)             | HR   | Patients with diabetes: annual review in the previous 12 months,        | Patients with diabetes: with blood pressure under control, with blood   |
|        |                                    |      | receiving antihypertensive medication, if needed, under regular        | glucose under control, with blood cholesterol under control             |
|        |                                    |      | treatment with lipid-lowering medication, if needed                     |                                                                         |
| 23     | Campanella et al. (2017)           | HR   |                                                                          | 30-day risk-adjusted mortality: AMI, CHF, PN                            |
| 23     | De et al. (2012)                   | HR   |                                                                          | Female life expectancy at birth; under-five mortality rates 0–5         |
| 23     | Serván-Mori et al. (2020)          | HR   | Antenatal care during pregnancy; opportune antenatal care; frequent    |                                                                         |
|        |                                    |      | institutionalized delivery; skilled services during childbirth;        |                                                                         |
|        |                                    |      | adequate maternal care                                                 |                                                                         |
| 30     | Ferreira et al. (2013)             | HR   | (Healthcare personnel; healthcare facilities)*                         | (Mortality rates)*                                                      |
| 30     | Gearhart and Michieka (2020)       | HR   |                                                                          | % births that are non-low birthweight; PYLG per 100,000 population      |
| 12     | Shimshak et al. (2007)             | NC/AC|                                                                          | Residents with: indwelling or external catheter, pressure sores/sclers, restraints |
| 13     | Bott et al. (2007)                 | NC/AC| Quality of care planning process (total deficiencies)                  |                                                                         |
| SubCAT | References                  | DMU     | Stage 1                                                                 | Outcome                                                                 |
|--------|-----------------------------|---------|-------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 13     | Min et al. (2016)           | NC/AC   | % long-stay residents: with daily pain, ADL decline, pressure ulcers; % residents: with restraints, urinary tract infection, falls in previous 30 days |
| 13     | Shen et al. (2020)          | NC/AC   | Human resource management                                               | Human resource management; rehabilitation quality management; problem-solving ability |
| 13     | Shwartz et al. (2016)       | NC/AC   | Risk-adjusted rates: pressure ulcers, functional decline, behavioral decline, mortality |
| 14     | Duffy et al. (2006)         | NC/AC   | % non-decubitus patients                                                | State inspection scores; decubiti-free days; potential decubiti days    |
| 14     | Kleinsorge and Karney (1992)| NC/AC   | State inspection scores; FTE employees per resident                     | State inspection scores                                                 |
| 15     | Avkiran and McCrystal (2014)| NC/AC   | Avg length of service of RNs                                             | ALOS of residents; hospitalizations; avg severity of hospitalizations; mortality rates; residents transferred |
| 20     | Kooreman (1994)             | NC/AC   |                                                                          |                                                                         |
| 20     | Lee et al. (2009)           | NC/AC   |                                                                          |                                                                         |
| 20     | Lin et al. (2017)           | NC/AC   |                                                                          |                                                                         |
| SubCAT | References | DMU | Stage 1 | Outcome |
|--------|------------|-----|---------|---------|
|        |            |     | Structural | Process | Outcome |
| 20     | Ni Luasa et al. (2018) | NC/AC | | | Residential charges |
| 21     | Garavaglia et al. (2011) | NC/AC | Extra nursing hours | | Unhandled complaint rates; incidence of major accidents |
| 21     | Zhang et al. (2019) | NC/AC | | | |
| 23     | Wichmann et al. (2018) | NC/AC | | | Quality of dying with dementia; quality of dying in long-term care; EURO-QOL 5D5L utilities |
| 24     | Zhang et al. (2008) | NC/AC | Acuity- and quality-adjusted: skilled nursing resident days, intermediate nursing resident days, other long-term care resident days | Acuity- and quality-adjusted: skilled nursing resident days, intermediate nursing resident days, other long-term care resident days | |
| 13     | Villamil et al. (2017) | OTHER (INS) | | Avg number of readmissions per patient with at least one emergency readmission post-surgery | Avg number of days per patient between the time of diagnosis and first treatment, of diagnostic studies before the first cancer diagnosis per patient, of histological studies per patient |
| 20     | Bryce et al. (2000) | OTHER (INS) | | | |
| 15     | Ghafari Someh et al. (2020) | OTHER (LAB) | Total scores of lab standards | Total scores of lab standards; avg waiting time for sampling; avg sample transfer time; test response time; correct tests; false tests | |
**Table 6** (continued)

| SubCAT | References                  | DMU          | Stage 1                                                                 |
|--------|-----------------------------|--------------|-------------------------------------------------------------------------|
|        |                             |              | Structural | Process                                                                 |
|        |                             |              | Outcome                                                               |
| 20     | Lamovšek and Klun (2020)    | OTHER (LAB)  | Support from diabetes care; access to diabetes care; continuity in     |
|        |                             |              | diabetes care; medical devices and medical treatment; diet and         |
|        |                             |              | exercise;                                                             |
| 15     | Borg et al. (2019)          | OTHER (PAT)  | Manage your diabetes; mood and energy; free of worries; not limited    |
|        |                             |              | by diabetes; not limited by blood sugar; general wellbeing; rates of   |
|        |                             |              | complaint handling; HbA1c; systolic blood pressure; LDL cholesterol    |
| 13     | Dowd et al. (2014)          | OTHER (PHY)  | Avoidable ED visits per beneficiary; ambulatory care sensitive         |
|        |                             |              | admissions per admission; claims-computable process quality measures    |
|        |                             |              | per eligible beneficiary;                                             |
|        |                             |              | Potentially preventable re-hospitalizations per admission              |
| 13     | Fiallos et al. (2017)       | OTHER (PHY)  | Rates of non-return patients within 72 h                               |
| 11     | Aduda et al. (2015)         | PHC          | Avg facility service quality index                                     |
| 11     | Milliken et al. (2011)      | PHC          | Quality-adjusted patients; performance-adjusted patients              |
### Table 6 (continued)

| SubCAT | References | DMU | Stage 1 |
|--------|------------|-----|---------|
|        |            |     | Structural | Process | Outcome |
| 11     | Ruiz-Rodriguez et al. (2016) | PHC | % adherence to national ANC guidelines; % women with at least one cytology over previous three years; % women 15–44 with at least one family planning counseling session at a health center |
| 12     | Glenngård (2013) | PHC | % patients satisfied with: continuity: doctor visits, nurse visits, consideration and respect, doctor visits, nurse visits; participation in decisions: doctor visits, nurse visits, adequate information: doctor visits, nurse visits; % patients recommending: doctors, nurses |
| 12     | McGarvey et al. (2019) | PHC | % prenatal patients who receive care in the first trimester | % low birth weight births |
| 13     | Rocha et al. (2014) | PHC | Full team; % non-precarious bonds in the team; doctor’s residence time; % clinics with graduated professionals; population under responsibility per population recommended; % professionals taking an introductory course; total hours of team training; % structure items available for each team | % scheduled queries among team queries; % actions developed |
| SubCAT | References                  | DMU     | Structural       | Process                  | Outcome                                      |
|--------|----------------------------|---------|------------------|--------------------------|----------------------------------------------|
|        |                            |         | Stage 1          |                          |                                              |
| 14     | Ferrera et al. (2014)      | PHC     | Quality index    | Quality index            |                                              |
| 14     | Gouveia et al. (2016)      | PHC     |                  |                          | Satisfaction with service quality           |
| 14     | Razzaq et al. (2013)       | PHC     |                  |                          | Patient satisfaction                         |
| 15     | Yuan et al. (2020)         | PHC     |                  |                          | Tangibility scores; assurance scores;        |
|        |                            |         |                  |                          | responsiveness scores; reliability scores;  |
|        |                            |         |                  |                          | empathy scores                               |
| 23     | Adang et al. (2016)        | PHC     | Performance indicator (stating and anticoagulants prescribing) |                          | LDL cholesterol; systolic blood pressure    |
|        |                            |         |                  |                          |                                              |
|        |                            |         | Stage 2          |                          |                                              |
| 12     | Kočišová et al. (2020)    | COUN    |                  |                          |                                              |
| 12     | Lo Storto and Gonchanuk (2017) | COUN    |                  |                          |                                              |
| 13     | Adang and Borm (2007)      | COUN    |                  |                          |                                              |
| 13     | Althin et al. (2019)       | COUN    |                  |                          |                                              |
| 13     | Kozuń-Cieślak (2020)       | COUN    |                  |                          |                                              |
| 13     | Ortega et al. (2017)       | COUN    |                  |                          |                                              |
| 13     | Santos et al. (2012)       | COUN    |                  |                          |                                              |
| 15     | Kim et al. (2020)          | COUN    |                  |                          |                                              |
| SubCAT | References | Stage 2 | DMU | Structural | Process | Outcome |
|--------|------------|---------|-----|------------|---------|---------|
| 23     | Puertas et al. (2020) | COUN    |     |            |         |         |
| 11     | Arrieta and Guillén (2017) | HC      |     |            |         |         |
| 11     | Azadeh et al. (2016) | HC      |     |            |         |         |
| 11     | Bilsel and Davutyan (2014) | HC      |     |            |         |         |
| 11     | Chang et al. (2011) | HC      |     |            |         |         |
| 11     | Chen et al. (2016) | HC      |     |            |         |         |
| 11     | Chen et al. (2019) | HC      |     |            |         |         |
| 11     | Du et al. (2014) | HC      |     |            |         |         |
| 11     | Giménez et al. (2019) | HC      |     |            |         |         |
| 11     | Guo et al. (2017) | HC      |     |            |         |         |
| 11     | Hu et al. (2012) | HC      |     |            |         |         |
| 11     | Kalis and Stracova (2019) | HC      |     |            |         |         |
| 11     | Matranga and Sapienza (2015) | HC      |     |            |         |         |
| 11     | Miller et al. (2017) | HC      |     |            |         |         |
| 11     | Morey et al. (1995) | HC      |     |            |         |         |
| 11     | Nedelea and Fannin (2017) | HC      |     |            |         |         |
| 11     | Park et al. (2018) | HC      |     |            |         |         |
| 11     | Sinuany-Stern et al. (2016) | HC      |     |            |         |         |
| 11     | Thanassoulis et al. (2016) | HC      |     |            |         |         |
| 11     | Wu et al. (2013) | HC      |     |            |         |         |

Perceived health status
| SubCAT | References                        | DMU  | Stage 2                                      |
|--------|-----------------------------------|------|---------------------------------------------|
|        |                                   |      | Structural Process Outcome                  |
| 13     | Chang and Lan (2010)              | HC   |                                             |
| 13     | Dismuke and Sena (2001)           | HC   |                                             |
| 13     | Kholghabad et al. (2019)          | HC   |                                             |
| 13     | Williams et al. (2016)            | HC   |                                             |
| 14     | Ferrier and Trivitt (2013)        | HC   |                                             |
| 14     | Karagiannis and Velentzas (2012)  | HC   |                                             |
| 14     | Navarro-Espigares and Torres (2011) | HC  |                                             |
| 14     | Nayar and Ozcan (2008)            | HC   |                                             |
| 14     | Prior (2006)                      | HC   |                                             |
| 14     | Sola and Prior (2001)             | HC   |                                             |
| 14     | Tiemann and Schreyögg (2009)      | HC   |                                             |
| 14     | van Ineveld et al. (2016)         | HC   |                                             |
| 14     | Yang and Zeng (2014)              | HC   |                                             |
| 15     | Cohen-Kadosh and Simuany-Stern (2020) | HC |                                             |
| 20     | Chu et al. (2002)                 | HC   | Teaching; TQM                               |
| 20     | Ferreira et al. (2020)            | HC   | Appropriate and timely care; access to care|
| 20     | Ferrier and Valdmanis (1996)      | HC   | Clinical safety; appropriate and timely care|
|        |                                   |      | Risk-adjusted mortality rates               |
### Table 6 (continued)

| SubCAT | References                   | DMU  | Stage 2                      | Outcome                                                   |
|--------|------------------------------|------|------------------------------|-----------------------------------------------------------|
| 20     | Gautam et al. (2013)         | HC   | Structural                   | Avg 30-day readmission rates: PN, HF                      |
| 20     | Gholami et al. (2015)        | HC   | Process                      | Risk-adjusted mortality index; risk-adjusted complications index; unadjusted mortality rates; unadjusted complication rates |
| 20     | Gok and Sezen (2013)         | HC   | Quality index                | Patient satisfaction                                      |
| 20     | Harrison and Coppola (2007)  | HC   | Hospital accreditation score; RN hours per inpatient day | Hospital accreditation score                             |
| 20     | Huerta et al. (2011)         | HC   | Use of computerized physician order entry | Hospital accreditation score                             |
| SubCAT | References | DMU | Structural | Process | Outcome |
|--------|------------|-----|------------|---------|---------|
| 20     | Kittelsen et al. (2015) | HC  | Teaching status |         | Patients readmitted acutely to inpatient care within 30 days of discharge; patients readmitted to inpatient care within 30 days of discharge and at least two days after discharge; out of hospital mortality from any cause (within 30, 90, 180, and 365 days after admission); pulmonary embolism/deep vein thrombosis; sepsis; accidental cut, puncture or hemorrhage during medical care; obstetric trauma; bedsores |
| 20     | Lindlbauer et al. (2016) | HC  | Certification  | Certification | Certification |
| 20     | Mujasi et al. (2016)     | HC  | Teaching status |         |         |
| 20     | Pai et al. (2019)        | HC  | Teaching status; FT RN per bed |         | Risk-adjusted readmission index; risk-adjusted mortality index |
| 21     | Barpanda and Sreekumar (2020) | HC  |               |         | Patient satisfaction |
| 21     | Clement et al. (2008)    | HC  | Teaching (major, minor, none) |         |         |
Table 6 (continued)

| SubCAT | References                | DMU   | Stage 2                                                                 |
|--------|---------------------------|-------|-------------------------------------------------------------------------|
|        |                           |       | Structural                  | Process                  | Outcome                                      |
| 21     | Dutta et al. (2014)       | HC    | ALOS; outpatients per inpatient day                                    |                         |                                             |
| 21     | Leleu et al. (2018)       | HC    | Teaching status; high technology status                                |                         |                                             |
| 21     | Lin et al. (2019)         | HC    | Accreditation               | Accreditation            | Accreditation                                |
| 21     | Morey et al. (1992)       | HC    | Member of Council of Teaching Hospitals; minor teaching hospital status; index of high-technology services |                         | Ratios of actual to forecasted deaths        |
| 21     | Mutter et al. (2010)      | HC    | Member of Council of Teaching Hospitals; minor teaching hospital status; index of high-technology services |                         |                                             |
| 21     | Nayar et al. (2013)       | HC    | Accreditation; teaching status                                         | Accreditation            | Accreditation                                |
| 21     | Valdmanis et al. (2008)   | HC    | Index of high-tech services; teaching status                           |                         |                                             |
| 21     | Wang et al. (2018)        | HC    | Poor overall quality of care                                           |                         | Poor overall quality of care                |
| 22     | Roth et al. (2019)        | HC    | Teaching status; electronic health records; technology service index   |                         |                                             |
| 24     | Chiu et al. (2008)        | HC    | High-tech equipment; licensed staff                                    |                         | Risk-adjusted mortality ratio: DRG level; level of major diagnoses |
| 24     | Li (2015)                 | HC    | Teaching status                                                         |                         |                                             |
| SubCAT | References                              | DMU | Stage 2                        | Process                                                                 | Outcome                                                                 |
|--------|----------------------------------------|-----|--------------------------------|------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 25     | Khushalani and Ozcan (2017)            | HC  | Teaching status               | (Hospital beds per 1,000 inhabitants; FTE nurses per 1,000 standard patients; FTE doctors per 1,000 standard patients; inpatient bed occupancy rates; OR capacity utilization)* | (Outpatient surgeries on potential outpatient procedures; waiting time before surgery; rates of first medical appointment within time; rates of surgeries within time; hip fracture surgery in first 48 h)* |
| 30     | Ferreira and Marques (2019)            | HC  |                               | (Hospital beds per 1,000 inhabitants; FTE nurses per 1,000 standard patients; FTE doctors per 1,000 standard patients; inpatient bed occupancy rates; OR capacity utilization)* | (Readmission rates within 30 days after discharge; rates of inpatients staying more than 30 days; decubitus ulcer rates; postoperative pulmonary embolism/deep vein thrombosis rates; postoperative sepsis rates; obstetric trauma rates (vaginal delivery with an instrument); obstetric trauma rates (vaginal delivery without instrument); in-hospital death rates for low severity cases; stillbirth rate; childhood mortality rates)* |
| 30     | Varabyova et al. (2017)                | HC  |                               | (Patient radiation exposure)*                                           | (Patient satisfaction; standardized mortality ratios)*                  |

* Denotes patient-oriented performance measures.
| Stage 2                          | SubCAT | References                  | DMU                |
|---------------------------------|--------|-----------------------------|--------------------|
|                                 | 20     | Laine et al. (2005)         | HC plus NC/AC       |
|                                 |        | Proportion of:             |                    |
|                                 |        | RNs; rooms with own toilet; |                    |
|                                 |        | single rooms               |                    |
|                                 |        | Prevalence of:             |                    |
|                                 |        | depression with no treatment, |                    |
|                                 |        | antipsychotic use, in the absence of |                    |
|                                 |        | psychotic and related conditions, |                    |
|                                 |        | anti-anxiety/hypnotic use, |                    |
|                                 |        | occasional or frequent bladder or bowel incontinence without a toileting plan, hypnotic use more than twice in last week; |                    |
|                                 |        | Lack of training/skill practice or range of motion for mobility dependent residents; little or no activity; use of 9 + different medications |                    |
|                                 | 14     | Obure et al. (2016)        | HC plus PHC        |
|                                 | 11     | Campoli et al. (2020)      | HR                 |
|                                 | 11     | Giuffrida and Gravelle (2001) | HR            |
|                                 | 11     | Karsak and Karadayi (2017) | HR                 |
|                                 | 11     | Thanassoulis et al. (1995) | HR                 |
|                                 |        | Prevalence of:             |                    |
|                                 |        | any injury; falls; behavioral symptoms affecting others; bladder or bowel incontinence; diagnosis or symptoms of depression; fecal impaction; urinary tract infections; weight loss; bedfast residents; indwelling catheters; daily physical restraints; stage 1–4 pressure ulcers; incidence of decline in late loss activities of daily living; new fractures; cognitive impairment; decline in the range of motion; new pressure ulcers |
| SubCAT | References | DMU  | Stage 2 |
|--------|------------|------|---------|
|        |            |      | Structural | Process | Outcome |
| 11     | Zheng et al. (2019) | HR   |           |         |         |
| 13     | Carrillo and Jorge (2017) | HR   |           |         |         |
| 13     | Ngobeni et al. (2020) | HR   |           |         |         |
| 13     | Pelone et al. (2012) | HR   |           |         |         |
| 13     | Salinas-Jiménez and Smith (1996) | HR   |           |         |         |
| 13     | Serván-Mori et al. (2018) | HR   |           |         |         |
| 13     | Takundwa et al. (2017) | HR   |           |         |         |
| 14     | Du (2018) | HR   |           |         |         |
| 14     | Polisena et al. (2010) | HR   |           |         |         |
| 23     | Allin et al. (2016) | HR   | Nursing hours per weighted inpatient cases |         | 30-day readmission rates |
| 23     | Amado and Dyson (2009) | HR   | Professional with training in diabetes was formal diabetes lead and delivered the majority of diabetes consultations; practice training status |         | Potentially avoidable discharges: HF, COPD; potentially avoidable long-term complications of diabetes |
| 23     | Campanella et al. (2017) | HR   |           |         |         |
| 23     | De et al. (2012) | HR   | Complete child immunization for children 12–23 months |         |         |
## Table 6 (continued)

| SubCAT | References                          | DMU | Stage 2                  | Outcome                                    |
|--------|-------------------------------------|-----|--------------------------|--------------------------------------------|
| 23     | Serván-Mori et al. (2020)           | HR  | Structural               | Maternal mortality ratios                  |
| 30     | Ferreira et al. (2013)              | HR  | Process                  | (Mortality rates)*; share of complaints     |
| 30     | Gearhart and Michieka (2020)        | HR  | Process                  | (Healthcare personnel; healthcare facilities)* |
| 12     | Shimshak et al. (2007)              | NC/AC|                          |                                            |
| 13     | Bott et al. (2007)                  | NC/AC|                          |                                            |
| 13     | Min et al. (2016)                   | NC/AC|                          |                                            |
| 13     | Shen et al. (2020)                  | NC/AC|                          |                                            |
| 13     | Schwartz et al. (2016)              | NC/AC|                          |                                            |
| 14     | Duffy et al. (2006)                 | NC/AC|                          |                                            |
| 14     | Kleinsorge and Karney (1992)        | NC/AC|                          |                                            |
| 15     | Avkiran and McCrystal (2014)        | NC/AC|                          |                                            |
| 20     | Kooreman (1994)                     | NC/AC| Presence of: daycare facilities, patient council; council of patients’ relatives; unrestricted visiting hours | Procedure for complaint handling; absence of physical therapy |
| 20     | Lee et al. (2009)                   | NC/AC| Number of deficiencies; director of nursing is the MDS nurse; the MDS nurse has a helper | Number of deficiencies |
| 20     | Lin et al. (2017)                   | NC/AC| Quality ratings          | Quality ratings                            |
| SubCAT | References | DMU       | Structural                                      | Process                                      | Outcome                                      |
|--------|------------|-----------|-------------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| 20     | Ni Luasa et al. (2018) | NC/AC     | Qualification (nurse has a diploma in gerontology) |                                               |                                               |
| 21     | Garavaglia et al. (2011) | NC/AC     | Residential quality                              | Residential quality                           |                                               |
| 21     | Zhang et al. (2019) | NC/AC     | Staff working time 2 + years; annual number of trainings; regular college graduates |                                               |                                               |
| 23     | Wichmann et al. (2018) | NC/AC     | Availability of palliative care and opioids      |                                               |                                               |
| 24     | Zhang et al. (2008) | NC/AC     | RN per total nursing staff; RN hours per resident day |                                               |                                               |
| 13     | Villamil et al. (2017) | OTHER (INS) | Federal qualification; state requires capital reserves; state guarantees for payment; state approves rates; consumer membership on the HMO board |                                               |                                               |
| 20     | Bryce et al. (2000) | OTHER (INS) |                                               |                                               |                                               |
| 15     | Ghafari Someh et al. (2020) | OTHER (LAB) |                                               |                                               |                                               |
| 20     | Lamovšek and Klun (2020) | OTHER (LAB) | Accreditation                                   | Accreditation                                 | Accreditation                                 |
| 15     | Borg et al. (2019) | OTHER (PAT) |                                               |                                               |                                               |
| 13     | Dowd et al. (2014) | OTHER (PHY) |                                               |                                               |                                               |
| 13     | Fiallos et al. (2017) | OTHER (PHY) |                                               |                                               |                                               |
| 11     | Aduda et al. (2015) | PHC       |                                               |                                               |                                               |
| SubCAT | References                  | DMU   | Stage 2                                |
|--------|-----------------------------|-------|----------------------------------------|
|        |                             |       | Structural               | Process | Outcome |
| 11     | Milliken et al. (2011)      | PHC   |                         |         |         |
| 11     | Ruiz-Rodriguez et al. (2016)| PHC   |                         |         |         |
| 12     | Glenngård (2013)            | PHC   |                         |         |         |
| 12     | McGarvey et al. (2019)      | PHC   |                         |         |         |
| 13     | Rocha et al. (2014)         | PHC   |                         |         |         |
| 14     | Ferrera et al. (2014)       | PHC   |                         |         |         |
| 14     | Gouveia et al. (2016)       | PHC   |                         |         |         |
| 14     | Razzaq et al. (2013)        | PHC   |                         |         |         |
| 15     | Yuan et al. (2020)          | PHC   |                         |         |         |
| 23     | Adang et al. (2016)         | PHC   | Training practice (GP education); nurse practicing CVRM tasks |         |         |

*Indicates variables used in conditional approaches
References

Adang E, Borm G (2007) Is there an association between economic performance and public satisfaction in health care? Eur J Health Econ 8:279–285
Adang E, Gerritsma A, Nouwens E, van Lieshout J, Wensing M (2016) Efficiency of the implementation of cardiovascular risk management in primary care practices: an observational study. Implement Sci 11:1–7
Aduda D, Ouma C, Onyango R, Onyango M, Bertrand J (2015) Voluntary medical male circumcision scale-up in Nyanza, Kenya: evaluating technical efficiency and productivity of service delivery. PLoS ONE 10:1–17
Ali A, Seiford L (1990) Translation invariance in data envelopment analysis. Oper Res Lett 9:403–405
Allin S, Grignon M, Wang L (2016) The determinants of efficiency in the Canadian health care system. Health Econ Policy Law 11:39–65
Allhin R et al (2019) Efficiency and productivity of cancer care in Europe. J Cancer Policy 21:100194
Amado C, dos Santos S (2009) Challenges for performance assessment and improvement in primary health care: the case of the Portuguese health centres. Health Policy 91:43–56
Amado C, Dyson R (2009) Exploring the use of DEA for formative evaluation in primary diabetes care: an application to compare English practices. J Oper Res Soc 60:1469–1482
Arrieta A, Guillén J (2017) Outpatient congestion leads to compromised care in Peruvian public hospital neonatal units. Health Care Manag Sci 20:157–164
Avkiran N, McCrystal A (2014) Intertemporal analysis of organizational productivity in residential aged care networks: scenario analyses for setting policy targets. Health Care Manag Sci 17:113–125
Azadeh A, Tohidii H, Zarrin M, Pashapour S, Moghaddam M (2016) An integrated algorithm for performance optimization of neurosurgical ICUs. Financial Account Manag 43:142–153
Bădin L, Daraio C, Simar L (2012) How to measure the impact of environmental factors in a nonparametric production model. Eur J Oper Res 223:818–833
Bădin L, Daraio C, Simar L (2014) Explaining inefficiency in nonparametric production models: the state of the art. Ann Oper Res 214:5–30
Banker R, Morey R (1986) Efficiency Analysis for Exogenously Fixed Inputs and Outputs. Oper Res 34:513–521
Barpanda S, Sreekumar N (2020) Performance analysis of hospitals in Kerala using data envelopment analysis model. J Health Manag 22:25–40
Bilsel M, Davutyan N (2014) Hospital efficiency with risk adjusted mortality as undesirable output: the Turkish case. Ann Oper Res 221:73–88
Borg S, Gerdtham U, Eeg-Olofsson K, Palaszewski B, Gudbjörnsdottir S (2019) Quality of life in chronic conditions using patient-reported measures and biomarkers: a DEA analysis in type 1 diabetes. Health Econ Rev 9:31
Bott M, Gojewski B, Lee R, Piamjariyakul U, Taunton R (2007) Care planning efficiency for nursing facilities. Nurs Econ 25:85–94
Bryce C, Engberg J, Wholey D (2000) Comparing the agreement among alternative models in evaluating HMO efficiency. Health Serv Res 35:509–528
Busse R, Panteli D, Quentin W (2019) An introduction to healthcare quality: defining and explaining its role in health systems. In: Busse R, Klazinga N, Panteli D, Quentin W (eds) Improving healthcare quality in Europe. Characteristics, effectiveness and implementation of different strategies. OECD, Paris, pp 3–18
Campanella P et al (2017) Hospital efficiency: how to spend less maintaining quality? Ann I Super Sanita 53:46–53
Campolli J, Alves Júnior P, Rossato F, Rebelatto D (2020) The efficiency of Bolsa Família Program to advance toward the Millennium Development Goals (MDGs): a human development indicator to Brazil. Socio Econ Plan Sci 71:100748
Carrillo M, Jorge J (2017) DEA-like efficiency ranking of regional health systems in Spain. Soc Indic Res 133:1133–1149
Chang L, Lan Y (2010) Has the national health insurance scheme improved hospital efficiency in Taiwan? Identifying factors that affects its efficiency. Afr J Bus Manag 4:3752–3760
Chang S, Hsiao H, Huang L, Chang H (2011) Taiwan quality indicator project and hospital productivity growth. Omega 39:14–22
Chen K, Chien L, Hsu Y, Yu M (2016) Metafrontier frameworks for studying hospital productivity growth and quality changes. Int J Qual Health Care 28:650–656
Chen K, Chen H, Chien L, Yu M (2019) Productivity growth and quality changes of hospitals in Taiwan: does ownership matter? Health Care Manag Sci 22:451–461
Chiu Y, Liu H, Chang M, Hsu L (2008) Use DEA model to study hospitals’ performance through medical quality. J Stat Manag Syst 11:823–846
Chu H, Liu S, Romeis J (2002) Does the implementation of responsibility centers, total quality management, and physician fee programs improve hospital efficiency? Evidence from Taiwan hospitals. Med Care 40:1223–1237
Clement J, Valdmanis V, Bazzoli G, Zhao M, Chukmaitov A (2008) Is more better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion. Health Care Manag Sci 11:67–77
Cohen-Kadosh S, Sinuany-Stern Z (2020) Hip fracture surgery efficiency in Israeli hospitals via a network data envelopment analysis. Cent Eur J Oper Res 28:251–277
Daraio C, Simar L (2007a) Advanced robust and nonparametric methods in efficiency analysis. Methodology and applications. Springer, New York
Daraio C, Simar L (2007b) Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. J Prod Anal 28:13–32
Daraio C, Simar L, Wilson P (2010) Testing whether two-stage estimation is meaningful in non-parametric models of production. LIDAM Discussion Papers ISBA 2010031, Leuven: Université catholique de Louvain
De P, Dhar A, Bhattacharya B (2012) Efficiency of health care system in India: an inter-state analysis using DEA approach. Soc Work Public Hlth 27:482–506
Dismuke C, Sena V (2001) Is there a trade-off between quality and productivity? The case of diagnostic technologies in Portugal. Ann Oper Res 107:101–116
Donabedian A (1980) Explorations in quality assessment and monitoring. Health Administration Press, Ann Arbor, Mich
Dowd B, Swenson T, Kane R, Parasharam S, Coulam R (2014) Can data envelopment analysis provide a scalar index of “value”? Health Econ 23:1465–1480
Du T (2018) Performance measurement of healthcare service and association discussion between quality and efficiency: evidence from 31 provinces of mainland China. Sustainability 10:74
Du J, Wang J, Chen Y, Chou S, Zhu J (2014) Incorporating health outcomes in Pennsylvania hospital efficiency: an additive super-efficiency DEA approach. Ann Oper Res 221:161–172
Duffy J, Fitzsimmons J, Jain N (2006) Identifying and studying “best-performing” services. Benchmarking 13:232–251
Dutta A, Bandyopadhyay S, Ghose A (2014) Measurement and determinants of public hospital efficiency in West Bengal, India. J Asian Public Policy 7:231–244
Ferreira D, Marques R (2019) Do quality and access to hospital services impact on their technical efficiency? Omega 86:218–236
Ferreira C, Marques R, Nicola P (2013) On evaluating health centers groups in Lisbon and Tagus Valley: efficiency, equity and quality. BMC Health Serv Res, 13
Ferreira D, Nunes A, Marques R (2020) Operational efficiency vs clinical safety, care appropriateness, timeliness, and access to health care. J Prod Anal 53:355–375
Ferrera J, Cebada E, Zamorano L (2014) The effect of quality and socio-demographic variables on efficiency measures in primary health care. Eur J Health Econ 15:289–302
Ferrier G, Trivitt J (2013) Incorporating quality into the measurement of hospital efficiency: a double DEA approach. J Prod Anal 40:337–355
Ferrier G, Valdmanis V (1996) Rural hospital performance and its correlates. J Prod Anal 7:63–80
Fiallos J, Patrick J, Michalowski W, Farion K (2017) Using data envelopment analysis for assessing the performance of pediatric emergency department physicians. Health Care Manag Sci 20:129–140
Garavaglia G, Lettieri E, Agasisti T, Lopez S (2011) Efficiency and quality of care in nursing homes: an Italian case study. Health Care Manag Sci 14:22–35
Gautam S, Hickef L, Johnson T, Mishra B (2013) Measuring the performance of Critical Access Hospitals in Missouri using data envelopment analysis. J Rural Health 29:150–158
Gearhart R, Michieka N (2020) A Non-Parametric Investigation of Supply Side Factors and Healthcare Efficiency in the US. J Prod Anal 54:59–74
Ghafari Someh N, Pishvae M, Sadjadi S, Soltani R (2020) Sustainable efficiency assessment of private diagnostic laboratories under uncertainty. J Model Manag 15:1069–1103
Gholami R, Hígón D, Emrouznejad A (2015) Hospital performance: Efficiency or quality? Can we have both with IT? Expert Syst Appl 42:5390–5400

Giménez V, Prieto W, Prior D, Tortosa-Ausina E (2019) Evaluation of efficiency in Colombian hospitals: An analysis for the post-reform period. Socio Econ Plan Sci 65:20–35

Giuffrida A, Gravelle H (2001) Measuring performance in primary care: econometric analysis and DEA. Appl Econ 33:163–175

Glenngård A (2013) Productivity and patient satisfaction in primary care—conflicting or compatible goals? Health Policy 111:157–165

Gok M, Sezen B (2013) Analyzing the ambiguous relationship between efficiency, quality and patient satisfaction in healthcare services: the case of public hospitals in Turkey. Health Policy 111:290–300

Golany B, Roll Y (1989) An application procedure for DEA. Omega 17:237–250

Gouveia M et al (2016) An application of value-based DEA to identify the best practices in primary health care. Or Spectrum 38:743–767

Guo H, Zhao Y, Niu T, Tsui K (2017) Hong Kong Hospital Authority resource efficiency evaluation: Via a novel DEA-Malmquist model and Tobit regression model. PloS One, 12

Halkos G, Petrou K (2019) Treating undesirable outputs in DEA: A critical review. Econ Anal Policy 62:97–104

Harrison J, Coppola M (2007) Is the quality of hospital care a function of leadership? Health Care Manag 26:263–272

Hollingsworth B, Dawson P, Maniadakis N (1999) Efficiency measurement of health care: a review of non-parametric methods and applications. Health Care Manag Sci 2:161–172

Hu H, Qi Q, Yang C (2012) Analysis of hospital technical efficiency in China: Effect of health insurance reform. China Econ Rev 23:865–877

Huerta T, Ford E, Ford W, Thompson M (2011) Realizing the Value Proposition: A Longitudinal Assessment of Hospitals’ Total Factor Productivity. J Healthc Eng 2:285–302

Hussey P, Wertheimer S, Mehrotra A (2013) The association between health care quality and cost: a systematic review. Ann Intern Med 158:27–34

Kalis R, Stracova E (2019) Using the DEA method to optimize the number of beds in the Slovak hospital sector. Econ Cas 67:725–742

Karagiannis R, Velentzas K (2012) Productivity and quality changes in Greek public hospitals. Oper Res 12:69–81

Karsak E, Karadayi M (2017) Imprecise DEA framework for evaluating health-care performance of districts. Kybernetes 46:706–727

Kholghabad H, Alisoltani N, Nazari-Shirouk S, Azadeh M, Moosakhani S (2019) A unique mathematical framework for optimizing patient satisfaction in emergency departments. Iran J Manage Stud 12:255–279

Khushalani J, Ozcan Y (2017) Are hospitals producing quality care efficiently? An analysis using Dynamic Network Data Envelopment Analysis (DEA). Socio Econ Plan Sci 60:15–23

Kim Y, ParkM, Atukeren E (2020) Healthcare and welfare policy efficiency in 34 developing countries in Asia. Int J Environ Res Public Health, 17

Kittelsen S et al (2015) Costs and Quality at the Hospital Level in the Nordic Countries. Health Econ 24(Suppl 2):140–163

Kleinsorge I, Karney D (1992) Management of nursing homes using data envelopment analysis. Socio Econ Plan Sci 26:57–71

Kočišová K, Cygańska M, Kludacz-Alessandri M (2020) The application of Data Envelopment Analysis for evaluation of efficiency of healthcare delivery for CVD patients. Econ Manag 23:96–113

Kohl S, Schoenfelder J, Fügener A, Brunner J (2018) The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals. Health Care Manag Sci 22:245–286

Kooreman P (1994) Nursing home care in the Netherlands: a nonparametric efficiency analysis. J Health Econ 13:301–316

Kozun-Cieślak G (2020) Is the efficiency of the healthcare system linked to the country’s economic performance? Beveridgeans versus Bismarckians. Acta Oeconomica 70:1–17

Laine J et al (2005) The association between quality of care and technical efficiency in long-term care. Int J Qual Health Care 17:259–267

Lamovšek N, Klun M (2020) Efficiency of medical laboratories after quality standard introduction: trend analysis of EU countries and case study from Slovenia. Cent Eur Public Admin Rev 18:143–163
Lee R, Bott M, Gajewski B, Taunton R (2009) Modeling efficiency at the process level: an examination of the care planning process in nursing homes. Health Serv Res 44:15–32

Leleu H, Al-Amin M, Rosko M, Valdmanis V (2018) A robust analysis of hospital efficiency and factors affecting variability. Health Serv Manage Res 31:33–42

Li C (2015) Hospital diagnostic aggregation and risk-adjusted quality. Health Serv Res 50:614–624

Lin J, Chen C, Peng T (2017) Study of the relevance of the quality of care, operating efficiency and inefficient quality competition of senior care facilities. Int J Env Res Pub Health 14:1–18

Lin F, Deng Y, Lu W, Kweh Q (2019) Impulse response function analysis of the impacts of hospital accreditations on hospital efficiency. Health Care Manag Sci 22:394–409

Lindlbauer I, Schreyögg J, Winter V (2016) Changes in technical efficiency after quality management certification: A DEA approach using difference-in-difference estimation with genetic matching in the hospital industry. Eur J Oper Res 250:1026–1036

Liu W, Sharp J (1999) DEA models via goal programming. In: Westermann G (ed) Data envelopment analysis in the public and private sector. Deutscher Universitäts-Verlag, Wiesbaden, pp 79–101

Liu W, Meng W, Li X, Zhang D (2010) DEA models with undesirable inputs and outputs. Ann Oper Res 173:177–194

Lo Storto C, Goncharuk A (2017) Efficiency vs effectiveness: a benchmarking study on European healthcare systems. Econ Sociol 10:102–115

Lovell C, Pastor J, Turner J (1987) Measuring macroeconomic performance in the OECD: A comparison of European and non-European countries. Eur J Oper Res 95:507–5018

Matranga D, Sapienza F (2015) Congestion analysis to evaluate the efficiency and appropriateness of hospitals in Sicily. Health Policy 119:324–332

McGarvey R, Thorsen A, Thorsen M, Reddy R (2019) Measuring efficiency of community health centers: a multi-model approach considering quality of care and heterogeneous operating environments. Health Care Manag Sci 22:489–511

Miller F, Wang J, Zhu J, Chen Y, Hockenberry J (2017) Investigation of the impact of the Massachusetts health care reform on hospital costs and quality of care. Ann Oper Res 250:129–146

Milliken O et al (2011) Comparative efficiency assessment of primary care service delivery models using data envelopment analysis. Can Public Pol 37:85–109

Min A, Park C, Scott L (2016) Evaluating technical efficiency of nursing care using data envelopment analysis and multilevel modeling. Western J Nurs Res 38:1489–1508

Morey R, Fine D, Loree S, Retzlaff-Roberts D, Tsubakitani S (1992) The trade-off between hospital cost and quality of care: an exploratory empirical analysis. Med Care 8:677–698

Morey R, Ozcan Y, Retzlaff-Roberts D, Fine D (1995) Estimating the hospital-wide cost differentials warranted for teaching hospitals: an alternative to regression approaches. Med Care 33:531–552

Mujasi P, Asbu E, Puig-Junoy J (2016) How efficient are referral hospitals in Uganda? A data envelopment analysis and tobit regression approach. BMC Health Serv Res, 16

Mutter R, Valdmanis V, Rosko M (2010) High versus lower quality hospitals: a comparison of environmental characteristics and technical efficiency. Health Serv Outcomes Res Method 10:134–153

Navarro-Espigares J, Torres E (2011) Efficiency and quality in health services: a crucial link. Serv Ind J 31:385–403

Nayar P, Ozcan Y (2008) Data envelopment analysis comparison of hospital efficiency and quality. J Med Syst 32:193–199

Nayar P, Ozcan Y, Yu F, Nguyen A (2013) Benchmarking urban acute care hospitals: efficiency and quality perspectives. Health Care Manage Rev 38:137–145

Nedelea I, Fannin J (2017) Testing for cost efficiency differences between two groups of rural hospitals. Int J Healthc Manag 10:57–65

Newhouse J (1970) Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital. Am Econ Rev 60:64–74

Ngobeni V, Breitenbach M, Aye G (2020) Technical efficiency of provincial public healthcare in South Africa. Cost Eff Resour Alloc 18:3

Ni Luasa S, Dineen D, Zieba M (2018) Technical and scale efficiency in public and private Irish nursing homes - a bootstrap DEA approach. Health Care Manag Sci 21:326–347

Obure C, Jacobs R, Guinness L, Mayhew S, Vassall A (2016) Does integration of HIV and sexual and reproductive health services improve technical efficiency in Kenya and Swaziland? An application of a two-stage semi parametric approach incorporating quality measures. Soc Sci Med 151:147–156
O’Neill L, Rauner M, Heidenberger K, Kraus M (2008) A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. Socioecon Plann Sci 42:158–189
Ortega B, Sanjuán J, Casquero A (2017) Determinants of efficiency in reducing child mortality in developing countries. The role of inequality and government effectiveness. Health Care Manag Sci 20:500–516
Pai D, Hosseini H, Brown R (2019) Does efficiency and quality of care affect hospital closures? Health Sys 8:17–30
Park J, Sung S, Kim Y, Lee B, Bae H (2018) A study on the quality-embedded efficiency measurement in DEA. Infor 56:247–263
Pastor J (1996) Translation invariance in data envelopment analysis. Ann Oper Res 66:93–102
Pelone F et al (2012) The measurement of relative efficiency of general practice and the implications for policy makers. Health Policy 107:258–268
Polisena M, Laporte A, Coye P, Croxford R (2010) Performance evaluation in home and community care. J Med Syst 34:291–297
Portela M et al (2016) Benchmarking hospitals through a web based platform. Benchmarking 23:722–739
Prior D (2006) Efficiency and total quality management in health care organizations: a dynamic frontier approach. Ann Oper Res 145:281–299
Pross C, Strumann C, Geissler A, Herwartz H, Klein N (2018) Quality and resource efficiency in hospital service provision: A geoaddititive stochastic frontier analysis of stroke quality of care in Germany. PloS One, 13
Puertas R, Marti L, Guaia-Martinez J (2020) Innovation, lifestyle, policy and socioeconomic factors: an analysis of European quality of life. Technol Forecast Soc 160:120209
Quentin W, Partanan V, Brownwood I, Klazinga N (2019) Measuring healthcare quality. In: Busse R, Klazinga N, Panteli D, Quentin W (eds) Improving healthcare quality in Europe. Characteristics, effectiveness and implementation of different strategies. OECD, Paris, pp 31–62
Razzaq S, Ali Chaudhary A, Razzaq Khan A (2013) Efficiency analysis of basic health units: a comparison of developed and deprived regions in Azad Jammu and Kashmir. Iran J Public Health 42:1223–1231
Rocha T, da Silva N, Barbosa A, Rodrigues J (2014) Human resource management in health and performance of work process in the primary health care - an efficiency analysis in a Brazilian municipality. J Health Manag 16:365–379
Roth A, Tucker A, Venkataraman S, Chilingerian J (2019) Being on the Productivity Frontier: Identifying “Triple Aim Performance” Hospitals. Prod Oper Manag 28:2165–2183
Ruiz-Rodriguez M, Rodriguez-Villamizar L, Heredia-Pi I (2016) Technical efficiency of women’s health prevention programs in Bucaramanga, Colombia: a four-stage analysis. BMC Health Serv Res 16:1–7
Salinas-Jiménez J, Smith P (1996) Data envelopment analysis applied to quality in primary health care. Ann Oper Res 67:141–161
Santos S, Amado C, Santos M (2012) Assessing the efficiency of mother-to-child HIV prevention in low- and middle-income countries using data envelopment analysis. Health Care Manag Sci 15:206–222
Scheel H (2001) Undesirable outputs in efficiency valuations. Eur J Oper Res 132:400–410
Serván-Mori E, Chivardi C, Mendoza M, Nigenda G (2018) A longitudinal assessment of technical efficiency in the outpatient production of maternal health services in México. Health Policy Plan 33:888–897
Serván-Mori E et al (2020) Tackling maternal mortality by improving technical efficiency in the production of primary health services: longitudinal evidence from the Mexican case. Health Care Manag Sci 23:571–584
Shen C, Hsu C, Lung F, Ly P (2020) Improving efficiency assessment of psychiatric halfway houses: a context-dependent data envelopment analysis approach. Healthcare 8:1–13
Shimshak D, Lenard M (2007) A two-model approach to measuring operating and quality efficiency with DEA. Infor 45:143–151
Shwartz M, Burgess J, Zhu J (2016) A DEA based composite measure of quality and its associated data uncertainty interval for health care provider profiling and pay-for-performance. Eur J Oper Res 253:489–502
Simar L, Wilson P (2007) Estimation and inference in two-stage, semi-parametric models of production processes. J Econ 136:31–64
Sinuany-Stern Z, Cohen-Kadosh S, Friedman L (2016) The relationship between the efficiency of orthopedic wards and the socio-economic status of their patients. Cent Eur J Oper Res 24:853–876
Sola M, Prior D (2001) Measuring productivity and quality changes using Data Envelopment Analysis: an application to Catalan hospitals. Financial Account Manag 17:219–245
Takundwa R, Jowett S, McLeod H, Peñaloza-Ramos M (2017) The effects of environmental factors on the efficiency of clinical commissioning groups in England: a data envelopment analysis. J Med Syst, 41
Thanassoulis E, Bousoffiane A, Dyson R (1995) Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. Eur J Oper Res 80:588–607
Thanassoulis E, Portela M, Graveney M (2016) Identifying the scope for savings at inpatient episode level: An illustration applying DEA to chronic obstructive pulmonary disease. Eur J Oper Res 255:570–582
Tiemann O, Schreyögg J (2009) Effects of ownership on hospital efficiency in Germany. Bus Res 2:115–145
Valdmanis V, Rosko M, Mutter R (2008) Hospital quality, efficiency, and input slack differentials. Health Serv Res 43:1830–1848
van Ineveld M, van Oostrum J, Vermeulen R, Steenhoek A, van de Klundert J (2016) Productivity and quality of Dutch hospitals during system reform. Health Care Manag Sci 19:279–290
Varabyova Y, Blankart C, Schreyögg J (2017) Using nonparametric conditional approach to integrate quality into efficiency analysis: empirical evidence from cardiology departments. Health Care Manag Sci 20:565–576
Villamil M et al (2017) Strategies for the quality assessment of the health care service providers in the treatment of gastric cancer in Colombia. BMC Health Serv Res 17:654
Wang L et al (2018) The determinants of the technical efficiency of acute inpatient care in Canada. Health Serv Res 53:4829–4847
Wichmann A et al (2018) Technical-efficiency analysis of end-of-life care in long-term care facilities within Europe: a cross-sectional study of deceased residents in 6 EU countries (PACE). PLoS ONE 13:e0204120
Williams C, Asi Y, Raffenaud A, Bagwell M, Zeini I (2016) The effect of information technology on hospital performance. Health Care Manag Sci 19:338–346
Wu C, Chang C, Chen P, Kuo K (2013) Efficiency and productivity change in Taiwan’s hospitals: a non-radial quality-adjusted measurement. Cent Eur J Oper Res 21:431–453
Yang J, Zeng W (2014) The trade-offs between efficiency and quality in the hospital production: some evidence from Shenzhen, China. China Econ Rev 31:166–184
Yuan H, Li H, Hou Z (2020) Is it worth outsourcing essential public health services in China? - Evidence from Beilin District of Xi’an. Int J Health Plann Mgmt 35:1486–1502
Zhang N, Unruh L, Wan T (2008) Has the Medicare prospective payment system led to increased nursing home efficiency? Health Serv Res 43:1043–1061
Zhang L, Zeng Y, Fang Y (2019) Evaluating the technical efficiency of care among long-term care facilities in Xiamen, China: based on data envelopment analysis and Tobit model. BMC Public Health 19:1230
Zheng D, Gong J, Zhang C (2019) Efficiency of medical service systems in the rural areas of Mainland China: a comparative study from 2013 to 2017. Public Health 171:139–147

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