ORIGINAL ARTICLE

CrossMark

Relative Technical Efficiency Assessment of Mental Health Services: A Systematic Review

Carlos R. García-Alonso[1] · Nerea Almeda[2] · José Alberto Salinas-Pérez[1] · Mencía R. Gutiérrez-Colosía[2] · Luis Salvador-Carulla[3]

Abstract

The current prevalence of mental disorders demands improved ways of the management and planning of mental health (MH) services. Relative technical efficiency (RTE) is an appropriate and robust indicator to support decision-making in health care, but it has not been applied significantly in MH. This article systematically reviews the empirical background of RTE in MH services following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Finally, 13 studies were included, and the findings provide new standard classifications of RTE variables, efficiency determinants and strategies to improve MH management and planning.

Keywords Systematic review · Relative technical efficiency · Mental health service · Mental health care · DEA

Introduction

The high prevalence of mental disorders has caused an enormous burden to society and incurred economic consequences worldwide (World Health Organization 2017). Taking into account this situation, developing strategies to improve the management of mental health (MH) care systems can be crucial. The World Health Organization (2005) has recognized the importance of designing appropriate MH policies to improve MH services; without them, the treatment of mental disorders is likely to be inefficient. Poorly informed decisions may be associated with effectiveness, efficiency and equity problems in health systems (Oxman et al. 2009). Thornicroft and Tansella (2009) established the nine fundamental principles to guide planning in community-based MH services. The ninth principle refers to the concept of "efficiency", which is described as the characteristic of the service to minimize inputs while maintaining constant the output production, or maximize outputs while maintaining constant the input consumed. Thus, the technical efficiency (TE) concept (Farrell 1957) is considered an appropriate indicator of service performance and has been used for planning purposes (Gattoufi et al. 2004). By assessing service TE, it is feasible to define their performance, to reach a better allocation of resources (number of health care institutions, beds or staff available) (Sun and Luo 2017) and to improve strategy design looking for a more reasonable distribution of total annual expenditures, balancing the number of patients assessed, hospitalized and treated, readjusting hospital sizes and improving networks with other health care services (Roh et al. 2013; Valdmanis et al. 2017). However, the identification of TE determinants (mostly environmental and structural factors, for example: the population size, prevalence of the disease, the socioeconomic status of the users, the service geographical location or their financial support and so on) (Leleu et al. 2017; Takundwa et al. 2017) and the differences in TE scores due to sample characteristics, for example comparing residential homes with health

✉ Nerea Almeda
nmalmeda@uloyola.es

1 Department of Quantitative Methods, Universidad Loyola Andalucía, Seville, Spain

2 Department of Psychology, Universidad Loyola Andalucía, Seville, Spain

3 Research School of Population Health, ANU College of Health and Medicine, Australian National University, Canberra, Australia

centres (Bjorkgren et al. 2001) are relevant for management purposes.

When TE is assessed in a group of comparable individuals, it becomes relative: relative TE (RTE; Torres-Jiménez et al. 2015). Data envelopment analyses (DEAs) is a nonparametric technique for RTE assessment of a set of comparable decision making units (DMUs; Charnes et al. 1978). Each DMU uses multiple inputs, such as beds and places, workforce capacity, availability of MH services and budget, to produce multiple outputs, such as number of admissions, average length of stay, discharges and readmissions. The original DEA model was the DEA-CCR (Charnes et al. 1978), which assumes a situation under constant returns to scale (CRS). This means that any change in the inputs (resources) results in a proportional change in the outputs. Later, Banker et al. (1984) developed a new DEA model called DEA-BCC, which is more flexible and assumes a variable returns to scale (VRS) situation. In this case, any change in the inputs does not involve a proportional change in the outputs. Both DEA techniques can focus their attention on the (i) inputs (input oriented DEA) or (ii) outputs (output oriented DEA). There are other models used to assess RTE, such as order-α and order-m methods (Torres-Jiménez et al. 2015), but they are used to a lesser extent. Numerous studies have assessed the RTE of health care services (e.g., Seiford 1997; Tavares 2002). Nevertheless, fewer studies have paid attention to specialized MH services.

There are various systematic reviews about the RTE assessment in health care services (Hussey et al. 2009; Kiadaliri et al. 2013; Pelone et al. 2015) but to date, there have been none for MH services. This paper aims to systematically review the empirical background of RTE in all types of MH services, and additionally, it offers a full review on the published evidence about how RTE is used in decision support.

Methods

For the systematic review process, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2009) was chosen. This systematic review was registered in the PROSPERO database (International Prospective Register of Systematic Reviews) with the number CRD42017056827.

Search Strategy

The search strategy was conducted on October 19, 2017. The final Boolean algorithm was initially piloted in the MEDLINE (Ovid version) database (Table 1). Later, it was driven in the PsycINFO, Web of Science, EMBASE, SCOPUS and ECONLIT databases. In addition, six other

databases were also explored in PROQUEST: Health and Medical Collection, Health Management Database, Nursing and Allied Health Database, Psychology Database, Public Health Database and Science Database. References from previous reviews on RTE in health care were also checked (Hollingsworth et al. 1999; Hussey et al. 2009; Kiadaliri et al. 2013; Pelone et al. 2015; Worthington 2004) as well as extensive DEA bibliographies (Emrouznejad et al. 2008; Emrouznejad and Yang 2017; Seiford 1996, 1997; Tavares 2002) to retrieve potential unidentified records.

The search strategy was developed by answering PICO fields:

- *Population* (*P*) all types of specialized MH services that deliver care for people with mental disorders. Due to the terminological variability of MH services, an extensive set of terms (Table 1) were included after the review of relevant publications (Johnson and Kuhlmann 2000; Killaspy et al. 2016; Priebe et al. 2009; Salvador-Carulla et al. 2013). This set comprised both MH care and general health care used for MH services. Considering that the care for substance use disorders is provided by different settings, for example: specialized treatment services for alcohol and drug abuse disorders, MH services, general health services and primary care among others (World Health Organization 2010), some specific terms in this field were also included in the search strategy in order to identify those services that provide MH care for these specific diagnoses.
- *Intervention* (*I*) the assessment of RTE by using DEA, order-α or order-m methods (Table 1). Additional terms were also used according to previous literature reviews of DEA (Emrouznejad et al. 2008; Emrouznejad and Yang 2017).
- *Comparison* (*C*) not applicable.
- *Outcome* (*O*) RTE scores (Table 1). Additional terms were also used according to previous literature reviews of DEA (Emrouznejad et al. 2008; Emrouznejad and Yang 2017).

Inclusion and Exclusion Criteria

Only the studies that assess RTE of specialized MH services for the population with a lived experience of mental disorders were included. "Specialized MH Care" typically include only part of the conditions listed as "mental disorders" in the International Classification of Diseases, 11th Revision (ICD-11) (World Health Organization 2018) and 5th Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association 2013). Services for alcohol and other drugs addictions are included in the MH care system in some jurisdictions while in others are provided by separate agencies. This is also the

Table 1 Search strategy piloted in MEDLINE: Ovid version

Identifies all types of mental health services [All fields (af)] and [Abstract (ab), Title (ti)]

1 "Mental health service*".af.

2 "Mental health care".ab,ti.

3 "Mental health system*".ab,ti.

4 "Inpatient mental health care".ab,ti.

5 "Psychiatric care".ab,ti.

6 "Psychiatric hospital*".ab,ti.

7 "Psychiatric outpatient*".ab,ti.

8 "Outpatient care".ab,ti.

9 "Addiction treatment unit".ab,ti.

10 "Substance abuse treatment center*".ab,ti.

11 "Community mental health cent*".ab,ti.

12 "Residential facilit*".ab,ti.

13 "Assisted living facilit*".ab,ti.

14 "Group home*".ab,ti.

15 "Halfway house*".ab,ti.

16 "Nursing home*".ab,ti.

17 "Residential mental health facilit*".ab,ti.

18 "Residential treatment".ab,ti.

19 "Residential care".ab,ti.

20 "Residential service*".ab,ti.

21 "Support* hous*".ab,ti.

22 "Support* home*".ab,ti.

23 "Support* accom*".ab,ti.

24 "Support* tenanc*".ab,ti.

25 "Shelter* house*".ab,ti.

26 "Shelter* home*".ab,ti.

27 "Assist* tenanc*".ab,ti.

28 "Floating support".ab,ti.

29 "Floating outreach".ab,ti.

30 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29

Identifies technical efficiency indicator [All fields] and [Abstract, Title]

31 "Relative technical efficiency".af.

32 "Technical efficiency".ab,ti.

33 "Organizational efficiency".af.

34 "Efficiency".af.

35 "Performance".ab,ti.

36 31 or 32 or 33 or 34 or 35

Identifies method for technical efficiency assessment [All fields]

37 "Data envelopment analysis".af.

38 "DEA".af.

39 "Decision making unit*".af.

40 "Linear programming".af.

41 "Mathematical model*".af.

42 "Operation* research".af.

43 "Order alpha".af.

44 "Order-α".af.

45 "Order-m".af.

46 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45

47 30 and 36 and 46

case for Neurodevelopmental Disorders. We follow in this review a broad approach to MH care and include all studies referring to "Mental Disorders" in the DSM-5, except for "Neurocognitive Disorders". Studies have to use DEA, order-α and/or order-m models and include RTE scores. Only peer-reviewed articles were included. Finally, there were no territorial, language or time constraints.

Quality Assessment

To assess the quality of the selected studies, we designed a specific checklist (Table 2) following those designed by (Hollingsworth 2008; Kiadaliri et al. 2013; Worthington 2004). Other relevant papers about RTE (Bahari and Emrouznejad 2014; Hatami-Marbini et al. 2011) and tools for the standardization of MH services (Johnson and Kuhlmann 2000; Salvador-Carulla et al. 2013) were also considered. The fields included in the quality assessment were as follows: use of any standardization system to classify MH services, sample, variable included (inputs and outputs), methodology, inclusion of expert's opinion, use of results for support decision making and, finally, the practical application of the results.

Results

Study Selection

The search strategy identified 2510 records. By checking previous systematic reviews and DEA bibliographies, six additional papers were found. After the duplicate removal, two paper authors (CG and NA) reviewed 1958 titles and abstracts, and 1817 records were excluded. The resulting 141 records were exhaustively full-text assessed for eligibility by the same two authors. One of the studies (Ferrier and Valdmanis 2002) was not available. Another author (JAS) resolved the disagreements independently. To assess the degree of agreement between the first two peer opinions, the intra-class correlation (ICC) analysis and Kappa index were used. The results showed that there was a very strong agreement (ICC = 0.885, α-Cronbach = 0.939, $p < 0.01$; and Kappa = 0.885, $p < 0.01$). Finally, 13 papers fulfilled inclusion criteria (Fig. 1).

Characteristics of the Selected Studies

Aims of the Studies

According to the taxonomy designed by Gattoufi et al. (2004), the common purposes of the selected papers are descriptive (service performance) and planning (support decision making). The demonstration that DEA is a suitable

and robust method for analysing RTE is another goal (Corredoira et al. 2011; Kontodimopoulos et al. 2006). Some studies (Gibert et al. 2010; Moran and Jacobs 2013; Nahra et al. 2009; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015) show a theoretical and empirical validation, introducing new methodological approaches. The identification of efficiency determinants (Alexander et al. 1998; Blank and Valdmanis 2005; Corredoira et al. 2011; Moran and; Jacobs 2013; Nahra et al. 2009) is also an objective. For further details, see Table 3.

Sample

The DMU used in the studies are health catchment areas, countries or a set of MH services. In the first case, the territory is divided into Small Health Areas (Gibert et al. 2010; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015). Moran and Jacobs (2013) analyse inpatient MH systems in European countries. The remaining papers analyse MH services (Alexander et al. 1998; Blank and Valdmanis 2005; Corredoira et al. 2011; Halsteinli et al. 2001; Kontodimopoulos et al. 2006; Nahra et al. 2009; Ozcan et al. 1996; Tyler et al. 1995; Yeh et al. 1997). For further details, see Table 3.

Type of MH Care Delivered (Standards)

Three studies (Gibert et al. 2010; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015) standardized the provision of care by using the European Service Mapping Schedule (Johnson and Kuhlmann 2000). Six studies analysed the RTE of services that provided only one main type of care that might be outpatient (Alexander et al. 1998; Corredoira et al. 2011; Halsteinli et al. 2001; Nahra et al. 2009) or inpatient care (Moran and Jacobs 2013; Ozcan et al. 1996). The remaining papers assessed more than one main type of care: residential, day and/or outpatient care (Blank and Valdmanis 2005; Gibert et al. 2010; Kontodimopoulos et al. 2006; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015; Tyler et al. 1995; Yeh et al. 1997). For further details, see Table 3.

Methods

All of the studies choose DEA for the RTE assessment, although one study (Torres-Jiménez et al. 2015) also uses the order-α and order-m methods. DEA can be used from an input and/or an output orientation. The majority select the first one (Alexander et al. 1998; Blank and Valdmanis 2005; Gibert et al. 2010; Kontodimopoulos et al. 2006; Moran and; Jacobs 2013; Nahra et al. 2009; Ozcan et al. 1996; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015; Tyler et al. 1995). DEA-CCR-CRS—is used in six studies (Alexander et al. 1998; Corredoira et al. 2011; Nahra et al. 2009; Ozcan

Table 2 Checklist for quality assessment

Quality assessment statements	1	2	3	4	5	6	7	8	9	10	11	12	13
Studies use any standardized system of classification of mental health services	X					X							
The sample is described in detail, including the main type of care provided		X	X	X	X	X	X	X	X	X	X	X	X
Studies assess more than one type of mental health care	X			X		X	X	X		X	X		X
The variables (inputs and outputs) chosen represent the mental health system	X	X	X	X	X	X	X	X	X	X	X	X	X
The outputs selected represent the mental health care outcomes	X		X	X	X	X			X	X	X	X	
Studies employ both input and output analysis													
Studies carry out both CRS and VRS							X	X	X	X	X	X	
Studies employ another method in addition to the DEA	X	X	X	X	X	X	X		X	X	X	X	X
Studies use sensitivity analysis	X	X	X	X	X	X	X	X	X	X	X	X	X
Studies assess the RTE in more than one scenario	X	X			X	X	X						
Studies assess the impact of efficiency determinants			X					X		X			
Studies take into consideration the relationship between expert opinion and methods applied	X			X		X							
Studies compare their results with other previous findings	X	X	X	X	X	X	X	X	X	X	X	X	X
Studies use results to support decision-making	X	X	X	X	X	X	X	X	X	X	X	X	X
The models have real-life applications	X		X		X	X	X	X	X	X	X	X	X

[1]Torres-Jiménez et al. (2015), [2]Moran and Jacobs (2013), [3]Corredoira et al. (2011), [4]Gibert et al. (2010), [5]Nahra et al. (2009), [6]Salvador-Carulla et al. (2007), [7]Kontodimopoulos et al. (2006), [8]Blank and Valdmanis (2005), [9]Halsteinli et al. (2001), [10]Alexander et al. (1998), [11]Yeh et al. (1997), [12]Ozcan et al. (1996), [13]Tyler et al. (1995)

Fig. 1 Flow chart of articles included and excluded after the systematic review
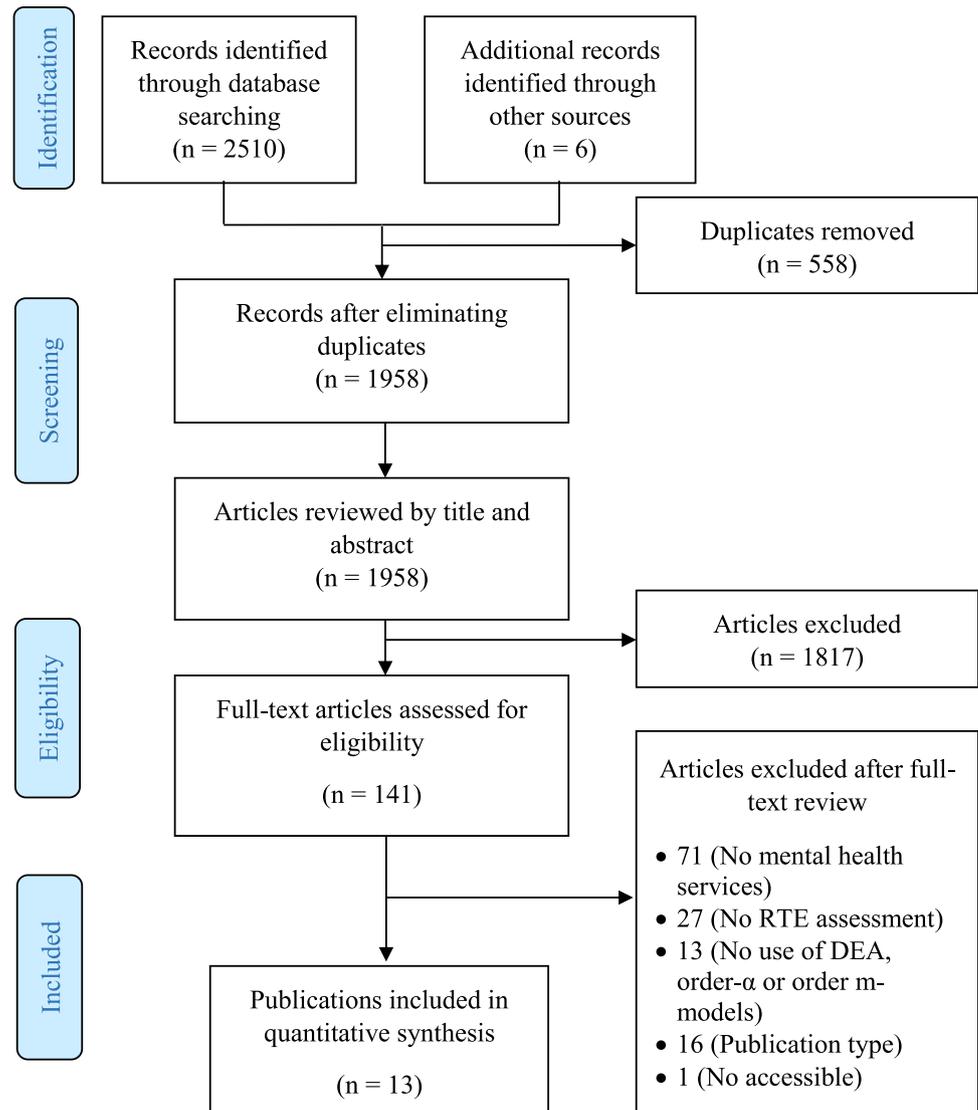
et al. 1996; Tyler et al. 1995; Yeh et al. 1997). DEA-BCC-VRS—is applied in five papers (Gibert et al. 2010; Halsteinli et al. 2001; Moran and Jacobs 2013; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015). Two papers (Blank and Valdmanis 2005; Kontodimopoulos et al. 2006) carry out both CRS and VRS.

RTE assessment is usually carried out in two steps. RTE scores are initially determined, and then their determinant factors are identified through regression analysis. Super-efficiency analysis and ordinary least squares regression (Nahra et al. 2009), the three stages RTE assessment (Blank and Valdmanis 2005) and bootstrapping (Moran and Jacobs 2013) can be considered DEA improvements. In Torres-Jiménez et al. (2015), DEA is integrated in a hybrid model with statistical simulation and artificial intelligence. The EbCA-DEA model (Gibert et al. 2010) allows the combination of expert judgements and DEA scores. DEA analysis is

also used to validate complex paradigms such as the operational model of basic MH community care (B-MHCC; Salvador-Carulla et al. 2007). In these cases, ICC and the Kappa index are used to determine the degree of agreement among experts' opinions and DEA scores. DEA scores are often analysed by standard statistical analysis such as Kolmogorov–Smirnov and Wilcoxon 2-sample tests. For further details, see Table 3.

Inputs and Outputs

DEA analyses in a comparative way the balance between inputs (mainly resources) and outputs. Due to the wide variability of inputs and outputs used, they were classified in categories according to the REFINEMENT Glossary of terms (Montagni et al. 2017). These categories are Group

Table 3 Characteristics of the selected studies: aims, sample, type of mental health care delivered and methods

Studies	Aims	Samples	Type of mental health care delivered	Methods
Torres-Jiménez et al. (2015)	Random RTE analysis	12 Small health-care areas in four regions in Spain	Acute and non-acute care (hospital), residential non-hospital care, day acute and non-acute care and other activities	Input oriented Monte Carlo BCC-DEA. Order-a and order-m methods; RTE statistical error; cluster analysis (k-means); ANOVA test; Tukey test and ICC
Moran and Jacobs (2013)	Cross-country mental healthcare systems comparison by assessing RTE	Inpatient mental healthcare systems in 32 OECD countries	Inpatient care	Bootstrapping DEA-VRS, input orientation; second stage-truncated regression; variance inflation factor and hierarchical agglomerative clustering
Corredoira et al. (2011)	RTE assessment to analyse performance factors	161 Addiction treatment clinics	Outpatient care	DEA-CRS with multivariate Tobit regression; output orientation
Gibert et al. (2010)	To introduce the EbCA-DEA model for supporting decision-making	12 Small Health Areas in 4 Autonomous Communities of Spain	Acute and non-acute care (hospital), residential non-hospital care, day acute and non-acute care and other activities	Expert-based cooperative analysis-DEA-BCC; input orientation; ICC and Kappa index
Nahra et al. (2009)	To introduce super-efficiency DEA and to analyse efficiency determinants on RTE	361 Substance abuse treatment units	Outpatient care	DEA-CCR with multivariate Tobit regression; super-efficiency DEA with ordinary least squares regression; input orientation
Salvador-Carulla et al. (2007)	To develop the B-MHCC and to utilize it to analyse RTE	12 Small Health Areas in 4 Autonomous Communities of Spain	Acute and non-acute care (hospital), residential non-hospital care, day acute and non-acute care and other activities	Input oriented BCC-DEA; Kolmogorov–Smirnov test; bivariate Pearson correlation; Spearman and Kendall's tau correlation; Delhi Expert Panel; ICC and Kappa index and Bayes predictive analysis
Kontodimopoulos et al. (2006)	To show a comparative efficiency assessment and to demonstrate the potentiality of DEA	90 Residential mental health facilities: 50 halfway houses, 8 nursing homes and 32 sheltered homes	Residential and day care (to the same users)	DEA-VRS and CRS; input orientation; independent sample T-test
Blank and Valdmanis (2005)	To analyse the cost inefficiency and the effect of the environment on allocative efficiency	71 Homes for mentally disabled	Residential and day care (to the same users)	DEA-CRS and VRS, input orientation; seemingly unrelated regression; test for misspecification error; Spearman's rank correlation test; T-test
Halsteinli et al. (2001)	To assess the productive efficiency in three policy issues	49 Clinics for children and youth	Outpatient care	DEA-VRS, output orientation; Kolmogorov–Smirnov statistic D+; ordinary T-test
Alexander et al. (1998)	To analyse the association between managed care participation and RTE and RTE determinants	442 Substance abuse treatment units	Outpatient care	DEA-CRS; input orientation; K-means rank sum test; descriptive and bivariate correlations; Probit equation; ordinary least square equation (Heckman estimation model)
Yeh et al. (1997)	To assess RTE by using DEA	40 Communities based youth services	Residential and day care	DEA-CRS, output orientation; variance analysis

Table 3 (continued)

Studies	Aims	Samples	Type of mental health care delivered	Methods
Ozcan et al. (1996)	To assess RTE by using DEA	85 Psychiatric hospitals	Inpatient care	DEA-CRS, input orientation; Wilcoxon 2-sample tests and T-tests
Tyler et al. (1995)	To analyse if Community Mental Health Centres show different RTE levels in the production of case management services	39 Community Mental Health Centres	Inpatient, day and outpatient care	DEA-CRS, input orientation; T-test; Wilcoxon z approximation

A: service provision, Group B: service utilization, Group C: quality of care and Group D: financing (Table 4).

Inputs and outputs in Group A are indicators related to the provision of MH care. Group B characterizes the service utilization. Group C refers to quality of care including the service availability, placement capacity (number of beds in residential care and places in day care), workforce capacity (grouping of all the professionals providing direct care and administrative staff) and effectiveness (achievement of intended outcomes by MH services). Group D includes financing items, and it is only used for inputs (see Table 4 for details).

Identification of Efficiency Determinants

Five studies analyse the variables that could affect RTE (Alexander et al. 1998; Blank and Valdmanis 2005; Corredoira et al. 2011; Moran and; Jacobs 2013; Nahra et al. 2009). Efficiency determinants are the independent variables in regression analysis. Taking into account the high variability, we have classified efficiency determinants by using the classification of Alexander et al. (1998), which organizes RTE determinants in four categories: (1) organizational, (2) environmental, (3) managed care and (4) financial. Nahra et al. (2009) also use this classification, except for the managed care field.

The first category involves service characteristics that can have a relevant impact on MH care provision, such as ownership, affiliation, patient characteristics and organizational size. Environmental factors refer to the response of services to the external pressure, e.g. number of providers and service localization. Managed care factors are associated with client treatments characteristics and the process for service provision. Finally, financial factors are related to economic resources (Table 5).

Reported Findings (a Meta-analysis Approximation)

Type of MH Care Delivered and RTE Scores

All the reviewed articles analyse MH services whose RTE is always assessed in a [0, 1] range (0 means absolute inefficiency and 1 means efficiency). When the RTE of MH services on average is available or is feasible to calculate, it varies between 0.37 and 0.97 (Table 6).

Regarding studies that analyse outpatient care, including clinics for children and youth and addiction treatment units, the RTE on average ($\overline{RTE}$) oscillates between 0.37 and 0.71. In the case of studies that assess inpatient care, $\overline{RTE}$ moves from 0.65 to 0.71.

Following with the studies that assess residential, day and/or outpatient care (two or more main types of care), $\overline{RTE}$ oscillates between [0.44, 0.97], being the lowest value

Table 4 New classification of inputs and outputs by using the Groups of the REFINEMENT Glossary of terms (Montagni et al. [2017])

Groups	Inputs of the studies	Outputs of the studies
A: service provision	*Facility* structured area (m²)[7]	*Mental health care* individual therapy-hours/week[5], group therapy-hours/week[5], family therapy-hours/week[5], individual therapy sessions/week[5], group therapy sessions/week[5], number of hours direct interventions[9], number of hours indirect interventions[9], individual therapy hours[10], group therapy hours[10], and family therapy hours[10]
B: service utilization	*Average length of stay*[2]	*Service utilization* acute hospital care utilization[1, 4 and 6], non-acute hospital care utilization[1, 4 and 6], community residential care utilization[1, 4 and 6], and day-hospital care plus other day-structured care excluding work and work-related services utilization[1, 4 and 6]
	Service user (i) severity of substance abuse disorder: number of high severity Level I patients[3], number of moderate severity Level I patients[3], and number of low severity Level I patients[3]	*Service user* clients receiving routine medical services on-site[5], clients receiving mental health services on-site[5], clients receiving employment services on-site[5], clients receiving financial services on-site[5], case management clients with serious mental illness[13], and case management clients without serious mental illness[13]
	(ii) Socio-demographic characteristics: youth population[11] and % youth in poverty[11]	*Bed occupancy* number of patients residing in each setting[7]
		Average length of stay corrected patients with slight disability days[8], corrected patients with moderate disability days[8], corrected patients with severe disability days[8]
		Hospital discharge discharges[2 and 12]
		Outpatient care average months in treatment[5] and outpatient visits[12]
C: quality of care	*Availability* acute hospital care availability[1 and 6]/types[4 and 6], non-acute hospital care availability[1 and 6], community residential care availability[1 and 6], and day-hospital care plus other day-structured care excluding work and work-related services availability[1, 4 and 6]	*Effectiveness* number of clients who completed the program[3] and number of clients with no drug use at discharge[3]
	Placement capacity acute hospital care beds[1 and 6], non-acute hospital care beds[1 and 6], community residential care beds[1 and 6], day-hospital care plus other day structured care excluding work and work-related services places[1 and 6], psychiatric beds[2 and 12], and number of operational hospital beds[12]	*Availability* residential services[11] and non-residential services[11]
	Workforce capacity: (i) level of studies: university educated staff[9] and college educated or administrative staff[9]	
	(ii) Staff specialty: psychiatrists[2], general personnel[8], nursing personnel[8], medical personnel[8], and auxiliary personnel[8]	
	(iii) Staff availability: full-time hours/week[5], part-time hours/week[5], consultant hours/per week[5], % treatment staff professionally trained[5], full-time staff[7], part-time staff[7], full-time staff hours[10], part-time staff hours[10], consultant hours[10], full-time personnel[12], full-time equivalent direct service staff[13], and full-time equivalent support staff[13]	
D: financing	*Diagnosis related group* number of diagnostic and special services (both inpatient and outpatient care)[12]	

Table 4 (continued)

Groups	Inputs of the studies	Outputs of the studies
	Global budget direct service budget[11], administration budget[11], normalized operating and capital expenditures[5], annual operational costs[7], regional prices[8]: general personnel, nursing personnel, medical personnel, auxiliary personnel and material supplies; amount of operational expenses[12], operating costs[13], and annual salaries[7]	

[1]Torres-Jiménez et al. (2015), [2]Moran and Jacobs (2013), [3]Corredoira et al. (2011), [4]Gibert et al. (2010), [5]Nahra et al. (2009), [6]Salvador-Carulla et al. (2007), [7]Kontodimopoulos et al. (2006), [8]Blank and Valdmanis (2005), [9]Halsteinli et al. (2001), [10]Alexander et al. (1998), [11]Yeh et al. (1997), [12]Ozcan et al. (1996), [13]Tyler et al. (1995)

for MH centres that deliver inpatient, day and outpatient care, while residential homes for mentally disabled achieved the highest one.

Characteristics of the Sample (DMU) and RTE Scores

Five articles classify the DMU according to specific characteristics (ownership, localization, size, delivered care and socio-economic status) in order to compare their differences in RTE scores. From an ownership perspective, relevant and significant differences from a statistical point of view have been found in private–public residential MH facilities (Kontodimopoulos et al. 2006) and in for-profit/non for profit psychiatric hospitals (Ozcan et al. 1996). According to localization and delivered care, there are significant differences between non-southern and southern psychiatric hospitals (Ozcan et al. 1996), and in centres that provided both inpatient and outpatient care versus those that deliver only outpatient care. On the other hand, localization, size and socioeconomic status characteristics do not show significant differences (Table 7).

Utilization of RTE Scores

All of the selected papers show the MH service performance by comparing it to a benchmark. Four studies focus their attention on improving traditional DEA techniques for analysing real situations. Here, methods such as Monte Carlo-DEA (Torres-Jiménez et al. 2015), EbCA-DEA (Gibert et al. 2010), B-MHCC (Salvador-Carulla et al. 2007) or super-efficiency (Nahra et al. 2009) can be highlighted. Gibert et al. (2010), Salvador-Carulla et al. (2007) and Torres-Jiménez et al. (2015) integrate these methods in decision support systems. In these papers, RTE is used for improving the system management. The majority of studies use RTE scores or averages as a guide to design real interventions, since they state that the RTE is a useful indicator in decision support.

Data Availability and RTE Scores

It is not possible to compare RTE scores among countries, areas and services because of data variability and data provision. For that reason, some authors (Blank and Valdmanis 2005; Corredoira et al. 2011; Gibert et al. 2010; Moran and; Jacobs 2013; Ozcan et al. 1996; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015; Tyler et al. 1995; Yeh et al. 1997) use proxy variables. This situation is worse when outcomes have to be measured (Halsteinli et al. 2001). Due to the sample sizes (MH units), sometimes it is required to increase the number of observations artificially (Torres-Jiménez et al. 2015). All of these factors affect the feasibility of obtaining robust conclusions for result generalization.

Table 5 Efficiency determinants classified in the categories defined by Alexander et al. (1998) and their impact on RTE in the selected studies

Studies	Categories based on Alexander et al. (1998)
Moran and Jacobs (2013)	*Organizational factors* alcohol consumption[NS], unemployment[NS], education as a proxy for socio-economic status[NS], and income as a proxy for socio-economic status[NS]
Corredoira et al. (2011)	*Organizational factors* county income per capita as a proxy for social support[NS], number of clinics per county square mile as a proxy for clinic accessibility↓*, and number of patients treated as a proxy for learning and experience[NS]
	Environmental factors state funding as a proxy for the management of market pressure on the clinic↓*
Nahra et al. (2009)	*Organizational factors* public ownership↓*, private non-for-profit ownership↓*, private for profit ownership[NS], mental health centre affiliation[NS], hospital affiliation[NS], other affiliation[NS], % Staff hour by volunteers↑*, organizational size↓*, age of unit in years[NS], experience of director at unit[NS], % clients prior treatment[NS], % clients unemployment[NS], % clients dual diagnoses[NS], % clients abuse multiple drugs[NS], and % clients African American↑*
	Environmental factors number of providers in county↑*, extent of non-price competition↓* and urban location dummy[NS]
	Managed care % clients in HMO or PPO[NS], % clients pre-authorization[NS] and % clients concurrent review[NS]
	Financial factors % clients private insurance↓*, % clients with Medicaid[NS] and % revenues by lump sum[NS]
Blank and Valdmanis (2005)	*Organizational factors* religious affiliation[NA], proportion of beds in socio-homes with at most eight beds[NA] and proportion of beds in socio-homes with more than eight beds[NA]
	Financial factors floor space as a proxy of capital resources[NA]
Alexander et al. (1998)	*Organizational factors* setting: unit is affiliated with a hospital↑*, setting: unit is affiliated with a mental health centre[NA], units provides methadone treatment↓*, ownership: unit is private for-profit[NA], ownership: unit is private not-for-profit[NA], unit has Joint Commission on the Accreditation of Healthcare Organizations Accreditation↓*, intensity of non-therapy services offered by unit[NA], clients severity: percentage clients African American[NA], clients severity: percentage clients with dual diagnoses[NA] and clients severity: percentage clients who abuse multiple drugs↓*
	Environmental factors level of non-price competition↓*, number of substance abuse providers in the county (competitors)[NA] and unit is in urban area location[NA]
	Managed care more than 10 clients in managed care arrangements (MCAs)[NS], percentage clients for whom MCAs specify treatments plans[NS], percentage clients for whom MCAs require written utilization review[NS], percentage clients for whom MCAs require correspondence with treatment team↑*, percentage clients for whom MCAs impose visit limits[NS], and percentage clients for whom MCAs impose sanctions[NS]
	Financial factors percentage clients with some insurance (public or private sources)↓* and percentage revenues in lump sum[NS]

NS no statistically significant in this specific study, *NA* no available

↑Significantly increases RTE at *($p \leq 0.05$) in this specific study

↓Significantly decreases RTE at *($p \leq 0.05$) in this specific study

Determinants of RTE

Five articles study the relationship among RTE scores and a number of determinants by using multivariate techniques. According to the classification of Alexander et al. (1998), the identified determinants are classified in four groups: organizational, environmental, managed care and financial (Table 5).

Sixteen variables show a statistically significant impact on MH services performance. Nine out of them can be classified within the organizational determinants, while four were included in the environmental group, one belongs to the managed care category and, finally, two were classified in the financial determinants. The impact of these variables depends on the situation analysed, that is: it varies according to the structure of the MH services, user characteristics, environmental socio-economic factors, etc. Therefore, a variable can have a positive or a negative impact on RTE in a specific situation and the contrary (or a neutral behaviour) in other (Table 5).

Strategies for Improving MH Care Management

The strategies for improving MH care management can be grouped as follows: financial, workforce capacity, placement capacity, service utilization, service provision, care management and research.

Regarding the first strategy, better resource management can improve MH service performance (Kontodimopoulos et al. 2006). Public MH units are more constrained by financial issues than private ones (Kontodimopoulos et al. 2006). Blank and Valdmanis (2005) state that an incorrect balance in resource allocation is the main cause of cost inefficiency. Therefore, both funders and providers should have tools for

Table 6 Type of mental health care delivered and relative technical efficiency (RTE) scores

Type of mental health care delivered (sample)	RTE scores on average ($\overline{RTE}$)	Studies
Outpatient care (clinics for children and youth)	0.71	Halsteinli et al. (2001)
Outpatient care (addiction treatment clinics)	0.71	Corredoira et al. (2011)
Outpatient care (substance abuse treatment units)	0.58	Nahra et al. (2009)
Outpatient care (substance abuse treatment units)	0.37	Alexander et al. (1998)
Inpatient care (inpatient mental healthcare systems)	0.71 ("efficient countries") 0.67 ("inefficient countries")	Moran and Jacobs (2013)
Inpatient care (psychiatric hospitals)	0.65	Ozcan et al. (1996)
Residential and day care (homes for mentally disabled)	0.97	Blank and Valdmanis (2005)
Acute hospital and day care (Small Health Areas)	0.87	Gibert et al. (2010)
Acute and non-acute hospital care, residential non-hospital care, day acute and non-acute care and other activities (Small Health Areas)	0.84	Torres-Jiménez et al. (2015)
Residential and day care (residential mental health facilities)	0.73 (VRS) 0.68 (CRS)	Kontodimopoulos et al. (2006)
Acute hospital care (Small Health Areas)	From 0.59 to 1	Salvador-Carulla et al. (2007)
Residential and day care (community based youth services)	0.6	Yeh et al. (1997)
Inpatient, day and outpatient care (Community Mental Health Centres)	0.44	Tyler et al. (1995)

VRS variable returns to scale, *CRS* constant returns to scale

Table 7 Characteristics of the sample (decision making units) and relative technical efficiency (RTE) scores

Characteristics of the sample (decision making units)		RTE scores on average ($\overline{RTE}$)	Studies
Ownership	Private residential MH facilities	0.87	Kontodimopoulos et al. (2006)
	Public residential MH facilities	0.69	
	For-profit hospitals	0.67	Ozcan et al. (1996)
	Non-profit hospitals	0.84	
Localization	Urban communities based youth services	0.55	Yeh et al. (1997)
	Rural communities based youth services	0.65	
	Non-southern psychiatric hospitals	0.80	Ozcan et al. (1996)
	Southern psychiatric hospitals	0.70	
Size	Larger MH clinics are more efficient than smaller ones	NA	Halsteinli et al. (2001)
	Small communities based youth services	0.55	Yeh et al. (1997)
	Large communities based youth services	0.69	
Delivered care	Centres that provided both inpatient and outpatient care	0.81	Tyler et al. (1995)
	Centres that deliver only outpatient care	0.47	
Socio-economic status	Rich communities based youth services	0.58	Yeh et al. (1997)
	Poor communities based youth services	0.64	

NA not available

estimating service performance (Corredoira et al. 2011). In addition, to improve efficiency, some authors propose reducing supply expenses (Ozcan et al. 1996), operating costs (Tyler et al. 1995) and the administrative service budget (Yeh et al. 1997).

Better workforce capacity management improves service performance. Consultative work in community, inpatient or primary care services, training time and the lack of labour specialization could be associated with lower levels of productivity (Halsteinli et al. 2001). Strategies to increase RTE are also focused on the reduction of placement capacity and/ or workforce capacity (Ozcan et al. 1996; Tyler et al. 1995).

Regarding service utilization, two papers state that the number of non-severe mental illness users (Tyler et al. 1995) and provided services (Yeh et al. 1997) should increase in any type of community MH services in order to improve the RTE scores. For another paper, the utilization of residential

care in the community should be increased everywhere (Salvador-Carulla et al. 2007). In contrast, this same paper does not find a consensus on whether it is better to increase or decrease the utilization of day care and acute hospital care services. Finally, the type of treated patients is also a relevant factor, so the RTE may improve when MH services attend patients suffering from non-severe mental illness because they require less resources (Tyler et al. 1995).

Findings related to service provision strategies indicate that the management of the provision of outpatient care is a more efficient way of treatment (Ozcan et al. 1996). In the care management group, there are no conclusive findings about how managed care affects RTE (Alexander et al. 1998), although this indicator is critical from a managerial point of view.

Finally, all studies highlight the relevance of research for improving MH care management. One study states that the inclusion of the B-MHCC in RTE assessment can help managers in their decision-making processes (Torres-Jiménez et al. 2015). Thus, the development of the B-MHCC model (Salvador-Carulla et al. 2007) provides a common framework for the improvement of resource allocation in MH services. Moreover, it is possible to decrease the uncertainty in decision-making processes, allowing for the design of evidence-based decisions, by using the EbCA-DEA model to introduce expert judgements in the analysis (Gibert et al. 2010). Finally, Blank and Valdmanis (2005) state that, by using the presented methodology, it is possible to identify where cost reductions can be carried out.

Quality of the Included Studies

The use of a standard system to classify MH services and MH care is not common. Three studies (Gibert et al. 2010; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015) belonging to the same research group use the European Service Mapping Schedule (Johnson and Kuhlmann 2000) codification system to classify MH services. Regarding the variables, inputs and outputs show the real situation of MH systems, and it is possible to understand how they use the former to produce the latter. It is frequent to use proxies due to the lack of output availability (e.g., discharges, service utilization and length of stay) (Corredoira et al. 2011; Gibert et al. 2010; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015).

No paper carries out both input- and output-oriented analyses within the same study. Two studies (Blank and Valdmanis 2005; Kontodimopoulos et al. 2006) use both CRS and VRS to assess RTE, and the latter makes some comparisons between them. Sensitivity analysis is common among all the studies, but only four (Kontodimopoulos et al. 2006; Moran and Jacobs 2013; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015) analyse more than one scenario

to obtain different perspectives of the real situation. Five studies (Alexander et al. 1998; Blank and Valdmanis 2005; Corredoira et al. 2011; Moran and; Jacobs 2013; Nahra et al. 2009) assess the efficiency determinants for determining the contextual factors that could have an impact on RTE. Three studies (Gibert et al. 2010; Salvador-Carulla et al. 2007; Torres-Jiménez et al. 2015) include expert opinions and compare them with DEA scores. All the study results aim to be used for improving MH care management and planning. Most of the papers analyse a real situation, except for those with a strong methodological focus (Gibert et al. 2010; Nahra et al. 2009) or when data are not available (Moran and Jacobs 2013). For further details, see Table 2.

Discussion

As far as we know, this is the first systematic review of RTE in MH services. The PRISMA guidelines (Moher et al. 2009) were followed by integrating a sensitive search strategy without restrictive filters. The search identified a large number of records that were filtered manually. This paper offers a comprehensive view of the current state of the art, including the collected findings and their potential use to design evidence-informed decisions in MH management and planning. The review follows the World Health Organization (2005) recommendations.

Although RTE assessment is widely applied in healthcare (e.g., Dulal 2017; Lin et al. 2017; van Ineveld et al. 2016), we found that it is not the case in MH care. Therefore, it is required that more studies analyse RTE in specialized MH services.

This paper highlights the large terminological variability related to the concept of "MH service" and the lack of use of standard classification tools, which makes it difficult to carry out feasible comparisons. For this reason, this paper focuses its attention on the main type of care delivered (Salvador-Carulla et al. 2013) instead of the name of the service. The findings highlight the importance of using standardized classification systems (Johnson and Kuhlmann 2000; Salvador-Carulla et al. 2013).

Inputs and outputs are classified in categories (Montagni et al. 2017) to provide clearer information because their diversity makes it difficult to achieve a global understanding and a meta-analysis. This paper provides the first approach for a common framework to categorise indicators for RTE analysis of MH services. RTE determinants are also classified (Alexander et al. 1998) in order to make the provision of information more understandable.

There is a common consensus that DEA is an appropriate and robust method to assess MH service performance. DEA results suggest ways to improve MH care management, reducing the potential waste of resources and/or

increasing the amount of outputs by identifying critical variables. DEA and its extensions (e.g., Monte Carlo-DEA, EbCA-DEA, B-MHCC, super-efficiency analyses and three stages DEA) can be integrated into more sophisticated tools to support evidence-informed decision-making, but the lack of available data is a barrier to apply the findings to real life.

RTE scores on average show that there is large variability among the selected MH services, regardless of the main type of care provided, and in the majority of cases, there is great room for improvement. When different MH services are compared in the same study, relevant differences are found. In the healthcare field, the association between the sample characteristics (size, ownership and location) and service performance brought significant differences (Bjorkgren et al. 2001).

RTE determinants play a relevant role in MH service performance. Findings in the health care field showed that hospital competition and ownership (Leleu et al. 2017), outpatient visits and average length of inpatient stay (Kirigia and Asbu 2013), population size, prevalence of disease and socio-economic status (unemployment rates) (Takundwa et al. 2017), and finally, size and location (Sikka et al. 2009) have an significant impact on RTE scores.

All of the strategies found agree that is necessary to conduct more research on MH service performance to integrate their findings into decision-making processes. The current high prevalence of mental disorders, the gap of unmet population needs, the costs associated with mental illness, and the pessimistic future estimations make it fundamental to develop new strategies for improving MH care management and provision. This paper shows the state of the art on the use of RTE in MH decision-making support by providing a full vision of inputs, outputs, RTE determinants and strategy categories to improve MH service management for providing better mental health care.

Compliance with Ethical Standards

Conflict of interest The authors declare no conflict of interest.

Ethical Approval This study does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent This study does not contain any studies with human participants.

References

References marked with an asterisk indicate studies included in the systematic review

*Alexander, J. A., Wheeler, J. R. C., Nahra, T. A., & Lemak, C. H. (1998). Managed care and technical efficiency in outpatient substance abuse treatment units. *Journal of Behavioral Health Services and Research, 25*(4), 377–396. https://doi.org/10.1007/BF02287509.

*Blank, J. L. T., & Valdmanis, V. G. (2005). A modified three-stage data envelopment analysis. *The European Journal of Health Economics, 6*(1), 65–72. https://doi.org/10.1007/s10198-004-0260-3.

*Corredoira, R. A., Chilingerian, J. A., & Kimberly, J. R. (2011). Analyzing performance in addiction treatment: An application of data envelopment analysis to the state of Maryland system. *Journal of Substance Abuse Treatment, 41*(1), 1–13. https://doi.org/10.1016/j.jsat.2011.01.006.

*Gibert, K., García-Alonso, C., & Salvador-Carulla, L. (2010). Integrating clinicians, knowledge and data: Expert-based cooperative analysis in healthcare decision support. *Health Research Policy and Systems, 8*, 28. https://doi.org/10.1186/1478-4505-8-28.

*Halsteinli, V., Kittelsen, S. A. C., & Magnussen, J. (2001). Scale, efficiency and organization in Norwegian psychiatric outpatient clinics for children. *The Journal of Mental Health Policy and Economics, 4*(2), 79–90.

*Kontodimopoulos, N., Bellali, T., Labiris, G., & Niakas, D. (2006). Investigating sources of inefficiency in residential mental health facilities. *Journal of Medical Systems, 30*(3), 169–176.

*Moran, V., & Jacobs, R. (2013). An international comparison of efficiency of inpatient mental health care systems. *Health Policy, 112*(1–2), 88–99. https://doi.org/10.1016/j.healthpol.2013.06.011.

*Nahra, T. A., Mendez, D., & Alexander, J. A. (2009). Employing super-efficiency analysis as an alternative to DEA: An application in outpatient substance abuse treatment. *European Journal of Operational Research, 196*(3), 1097–1106. https://doi.org/10.1016/j.ejor.2008.04.022.

*Ozcan, Y. A., McCue, M. J., & Okasha, A. A. (1996). Measuring the technical efficiency of psychiatric hospitals. *Journal of Medical Systems, 20*(3), 141–150. https://doi.org/10.1007/BF02281992.

*Salvador-Carulla, L., García-Alonso, C. R., González-Caballero, J. L., & Garrido-Cumbrera, M. (2007). Use of an operational model of community care to assess technical efficiency and benchmarking of small mental health areas in Spain. *The Journal of Mental Health Policy and Economics, 10*(2), 87–100.

*Torres-Jiménez, M., García-Alonso, C. R., Salvador-Carulla, L., & Fernández-Rodríguez, V. (2015). Evaluation of system efficiency using the Monte Carlo DEA: The case of small health areas. *European Journal of Operational Research, 242*(2), 525–535. https://doi.org/10.1016/j.ejor.2014.10.019.

*Tyler, L. H., Ozcan, Y. A., & Wogen, S. E. (1995). Mental health case management and technical efficiency. *Journal of Medical Systems, 19*(5), 413–423.

*Yeh, J., White, K. R., & Ozcan, Y. A. (1997). Efficiency evaluation of community-based youth services in Virginia. *Community Mental Health Journal, 33*(6), 487–499. https://doi.org/10.1023/A:1025096502578.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Arlington, TX: American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596.744053.

Bahari, A. R., & Emrouznejad, A. (2014). Influential DMUs and outlier detection in data envelopment analysis with an application

to health care. *Annals of Operations Research, 223*(1), 95–108. https://doi.org/10.1007/s10479-014-1604-8.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092. http://econpapers.repec.org/RePEc:inm:ormnsc:v:30:y:1984:i:9:p:1078-1092.

Bjorkgren, M. A., Hakkinen, U., & Linna, M. (2001). Measuring efficiency of long-term care units in Finland. *Health Care Management Science, 4*(3), 193–200.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444. http://econpapers.repec.org/RePEc:eee:ejores:v:2:y:1978:i:6:p:429-444.

Dulal, R. (2017). Technical efficiency of nursing homes: Do five-star quality ratings matter? *Health Care Management Science.* https://doi.org/10.1007/s10729-017-9392-8.

Emrouznejad, A., Parker, B. R., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-economic Planning Sciences, 42*(3), 151–157. https://doi.org/10.1016/j.seps.2007.07.002.

Emrouznejad, A., & Yang, G. (2017). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-economic Planning Sciences.* https://doi.org/10.1016/j.seps.2017.01.008.

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

Ferrier, G. D., & Valdmanis, V. G. (2002). Exploring psychiatric hospital performance using data envelopment analysis and cluster analysis. *Journal d'Economie Medicale, 20*(3–4), 143–153. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0036588661&partnerID=40&md5=4dfe8cbc35afb2b4514aab556a3d6e56.

Gattoufi, S., Oral, M., & Reisman, A. (2004). A taxonomy for data envelopment analysis. *Socio-economic Planning Sciences, 38*(2–3), 141–158. https://doi.org/10.1016/S0038-0121(03)00022-3.

Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making. *European Journal of Operational Research, 214*(3), 457–472. https://doi.org/10.1016/j.ejor.2011.02.001.

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

Hollingsworth, B., Dawson, P. J., & Maniadakis, N. (1999). Efficiency measurement of health care: A review of non-parametric methods and applications. *Health Care Management Science, 2*(3), 161–172.

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

Johnson, S., & Kuhlmann, R. (2000). The European Service Mapping Schedule (ESMS): Development of an instrument for the description and classification of mental health services. *Acta psychiatrica Scandinavica. Supplementum, 405*, 14–23.

Kiadaliri, A. A., Jafari, M., & Gerdtham, U.-G. (2013). Frontier-based techniques in measuring hospital efficiency in Iran: A systematic review and meta-regression analysis. *BMC Health Services Research, 13*(1), 312. https://doi.org/10.1186/1472-6963-13-312.

Killaspy, H., White, S., Dowling, S., Krotofil, J., McPherson, P., Sandhu, S., et al. (2016). Adaptation of the Quality Indicator for Rehabilitative Care (QuIRC) for use in mental health supported accommodation services (QuIRC-SA). *BMC Psychiatry, 16*(1), 101. https://doi.org/10.1186/s12888-016-0799-4.

Kirigia, J. M., & Asbu, E. Z. (2013). Technical and scale efficiency of public community hospitals in Eritrea: An exploratory study. *Health Economics Review, 3*(1), 6. https://doi.org/10.1186/2191-1991-3-6.

Leleu, H., Al-Amin, M., Rosko, M., & Valdmanis, V. G. (2017). A robust analysis of hospital efficiency and factors affecting variability. *Health Services Management Research.* https://doi.org/10.1177/0951484817730537.

Lin, J. R., Chen, C. Y., & Peng, T. K. (2017). Study of the relevance of the quality of care, operating efficiency and inefficient quality competition of senior care facilities. *International Journal of Environmental Research and Public Health.* https://doi.org/10.3390/ijerph14091047.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269.

Montagni, I., Salvador-Carulla, L., Mcdaid, D., Straßmayr, C., Endel, F., Näätänen, P., et al. (2017). The REFINEMENT Glossary of terms: An international terminology for mental health systems assessment. *Administration and Policy in Mental Health and Mental Health Services Research.* https://doi.org/10.1007/s10488-017-0826-x.

Oxman, A. D., Lavis, J. N., Lewin, S., & Fretheim, A. (2009). SUPPORT Tools for evidence-informed health Policymaking (STP)1: What is evidence-informed policymaking? *Health Research Policy and Systems, 7*, 1–7. https://doi.org/10.1186/1478-This.

Pelone, F., Kringos, D. S., Romaniello, A., Archibugi, M., Salsiri, C., & Ricciardi, W. (2015). Primary care efficiency measurement using data envelopment analysis: A systematic review. *Journal of Medical Systems.* https://doi.org/10.1007/s10916-014-0156-4.

Priebe, S., Saidi, M., Want, A., Mangalore, R., & Knapp, M. (2009). Housing services for people with mental disorders in England: Patient characteristics, care provision and costs. *Social Psychiatry and Psychiatric Epidemiology, 44*(10), 805–814. https://doi.org/10.1007/s00127-009-0001-0.

References marked with an asterisk indicate studies included in the systematic review.

Roh, C.-Y., Moon, M. J., & Jung, K. (2013). Efficiency disparities among community hospitals in Tennessee: Do size, location, ownership, and network matter? *Journal of Health Care for the Poor and Underserved, 24*(4), 1816–1833. https://doi.org/10.1353/hpu.2013.0175.

Salvador-Carulla, L., Alvarez-Galvez, J., Romero, C., Gutiérrez-Colosía, M. R., Weber, G., McDaid, D., et al. (2013). Evaluation of an integrated system for classification, assessment and comparison of services for long-term care in Europe: The eDESDE-LTC study. *BMC Health Services Research, 13*(1), 218. https://doi.org/10.1186/1472-6963-13-218.

Seiford, L. M. (1996). Data envelopment analysis: The evolution of the state of the art (1978–1995). *Journal of Productivity Analysis, 7*(2), 99–137. https://doi.org/10.1007/BF00157037.

Seiford, L. M. (1997). A bibliography for Data Envelopment Analysis (1978–1996). *Annals of Operations Research, 73*, 393–438. https://doi.org/10.1023/A:1018949800069.

Sikka, V., Luke, R. D., & Ozcan, Y. A. (2009). The efficiency of hospital-based clusters: Evaluating system performance using data envelopment analysis. *Health Care Management Review, 34*(3), 251–261. https://doi.org/10.1097/HMR.0b013e3181a16ba7.

Sun, J., & Luo, H. (2017). Evaluation on equality and efficiency of health resources allocation and health services utilization in China. *International Journal for Equity in Health, 16*(1), 127. https://doi.org/10.1186/s12939-017-0614-y.

Takundwa, R., Jowett, S., McLeod, H., & Peñaloza-Ramos, M. C. (2017). The effects of environmental factors on the efficiency of clinical commissioning groups in England: A data envelopment

analysis. *Journal of Medical Systems, 41*(6), 1–7. https://doi.org/10.1007/s10916-017-0740-5.

Tavares, G. (2002). *A bibliography of Data Envelopment Analysis (1978–2001)*. RRR (Rutgers Research Report, p 186). http://rutcor.rutgers.edu/2002.html, http://rutcor.rutgers.edu/pub/rrr/reports2002/1_2002.pdf.

Thornicroft, G., & Tansella, M. (2009). *Better mental health care* (Cambridge edition). Cambridge University Press.

Valdmanis, V. G., Rosko, M. D., Leleu, H., & Mukamel, D. B. (2017). Assessing overall, technical, and scale efficiency among home health care agencies. *Health Care Management Science, 20*(2), 265–275. https://doi.org/10.1007/s10729-015-9351-1.

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 Management Science, 19*(3), 279–290. https://doi.org/10.1007/s10729-015-9321-7.

World Health Organization. (2005). *Mental health policy, plans and programmes*. Mental health policy and service guidance packages (pp 1–122). http://www.who.int/mental_health/policy/services/2_policy plans prog_WEB_07.pdf.

World Health Organization. (2010). *ATLAS on Substance Use (2010)—Resources for the prevention and treatment of substance use disorders. Chapter 2. Health Services*. France. http://www.who.int/substance_abuse/activities/msbatlaschtwo.pdf?ua=1.

World Health Organization. (2017). *Mental disorders*. http://www.who.int/mediacentre/factsheets/fs396/en/.

World Health Organization. (2018). *The International Classification of Diseases 11th Revision*. Retrieved October 8, 2018, from https://icd.who.int/.

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