European Resaerch Studies Journal
Volume XXIV, Issue 2, 2021

pp. 1095-1117

Health System Efficiency in European Countries: Network Data Envelopment Analysis Approach

Submitted 09/04/21, 1st revision 22/04/21, 2nd revision 18/05/21, accepted 25/05/21

Justyna Kujawska[1]

Abstract:

Purpose: *The article's main aim is to investigate the effectiveness of health systems in European countries based on EUROSTAT data. A comparative analysis of the health systems' effectiveness in different countries is based on their improvement (reform), using the best practices approach.*
Design/Methodology/Approach: *The network DEA model and a slack-based model (NDEA – SBM) are used. A non-oriented model is used. The research sample covers 30 countries (28 EU plus Norway and Island). The health system considers two factors: lifestyle (LF) and primary medical care resources (MC). Lifestyle factors included, alcohol consumption, smoking, and being overweight. The primary resources of medical care are medical personnel, hospital beds, and finances. The subjective assessment of health status and healthy life expectancy is taken as direct outputs. As an intermediate product (link), the expenditure on prevention is assumed as a percentage of GDP.*
Findings: *Health systems in five countries are identified as fully efficient. These countries have also achieved total efficiency for both divisions, lifestyle factors and medical care. The average efficiency of health systems for all countries is low and amounts to 0.619, and the average efficiency for the LF division is 0.580, with huge variations between countries. In the MC division, the average efficiency for all countries is 0.72. However, the difference between countries is more minor. For inefficient countries, the projection of necessary changes to achieve total efficiency has been calculated.*
Practical Implications: *The network DEA model allows a better understanding of the functioning of the complex health care system by analyzing the effectiveness of two separate areas (lifestyle and health care). The values of the forecasted variables are also determined, inputs, outputs, and the linking variable that may help determine priority actions in the field of improving the efficiency of health systems.*
Originality/value: *This paper contributes to the literature by applying the network model to assess the effectiveness of health systems. It enables simultaneous research of various areas related to health. Few publications have attempted to use the network model in the area of health protection.*

Keywords: *Health system, lifestyle factors, efficiency, network data envelopment analysis.*

JEL classification: *C14, C61, H51,I12.*

Paper Type: *Research study.*

[1]*Gdańsk University of Technology, Faculty of Management and Economics*, Poland, e-mail: Justyna.Kujawska@pg.edu.pl;

1. Introduction

Many factors influence health outcomes, therefore, their inclusion inefficiency analyses are critical public health issues (Rettenmaier and Wang, 2013; Mocan and Altindag, 2014) and World Health Organization (WHO, 2000). The World Health Organization (WHO, 2000) states that it is difficult to define what a health system is, what it consists of, and where it begins. The health system is assumed to cover all activities aimed primarily at promoting, restoring, or maintaining the health of populations conducted with people, institutions, and resources organized according to agreed national policies. The Robert Wood Johnson Foundation stresses that "health is more than just healthcare" (RWJF, 2019). Therefore, the analysis of health systems cannot be limited to determining the technical efficiency of using the resources involved, such as medical personnel, medical care infrastructure, and financial resources. This set of resources is extended to cover non-medical factors affecting the health status of the population (OECD, 2010; RWJF, 2019; Rettenmaier and Wang, 2013). Woolf and Aron (2013) stress that to reflect the complexity of the health system entirely, the links between public health (population-based services) and medical care (provided to individual patients) must be considered. In their view, both components should be taken into account in international comparisons of health systems.

Worldwide, many comparative studies of health systems have been conducted across different groups of countries, in most cases using non-parametric data envelopment analysis (DEA). The approaches can be categorized into two main types based on the DEA model (Mitropoulos, 2019; Kao, 2014; Färe and Grosskopf, 2000; Ozcan and Khushalani, 2017). The traditional approach treats units (called decision-making units, DMUs) subject to assessment (health systems) as black boxes, assuming that the production process is a function of initial inputs and final outputs without information about the activities performed within each DMU, (Retzlaff-Roberts *et al.*, 2004; Cheng and Zervopoulos, 2014; González *et al.*, 2010; Hadad *et al.*, 2013; Mitropoulos, 2019). This approach makes it difficult to distinguish the efficiency of different components of a complex health system, such as public health and health care, and their interaction. To overcome these problems, network DEA models (NDEA) may have two or more divisions or stages that are interlinked (Mitropoulos, 2019; Ozcan and Khushalani, 2017), is increasingly common. Outputs of one division are considered intermediate products that simultaneously constitute the inputs of the following one.

As a result, the efficiency of both individual divisions and the entire complex system can be assessed. The NDEA model first appeared in 2000 when it was formalized by Färe and Grosskopf (2000) for the Swedish Institute for Health Economics, but NDEA remains a relatively rarely used tool in health system research. Based on an analysis of 262 articles on DEA applications in health care published between 2005 and 2016, with particular emphasis on hospitals, Kohl *et al.* (2019) showed that of 330 models used (in many articles, multiple models were evaluated), only two articles used an NDEA model, and the vast majority of publications used basic radial models (CCR and BCC in 78% of applications). This fact is also confirmed in studies by Cantor and Poh (2018), who reviewed 57 articles using DEA in the health care sector.

A comparative analysis of health systems and health care systems in different countries is instrumental, as it enables a better understanding of how such systems work. The results of such analyses may serve as a foundation for system improvement based on best practices (Mitropoulos, 2019; Papanicolas and Smith, 2013). In most countries, initiatives to improve the performance of health systems have been a critical policy issue for many years (OECD, 2010), as health spending is one of the main areas of public spending (de Cos and Moral-Benito, 2014; Yaya and Danhoundo, 2015; Afonso and Aubyn, 2005; Afonso and Aubyn, 2011). Mitropoulos (2019) stressed that increasing the efficiency of publicly funded health systems should ensure better access to services for the public and should not affect the quality in EU countries, health care spending grew faster than national income, which according to Medeiros and Schwierz (2015), is to a large extent a result of population aging, medical innovations, and the observed inefficiency of health care systems. They also stressed that health outcomes are influenced by past and present lifestyle behavior and environmental factors outside the immediate control of the health system.

This article compares the efficiency of health systems in selected European countries, considering essential health care resources and population lifestyle factors. This article contributes to the literature in various ways. The network model of DEA, rarely implemented in previous studies of health systems, is used. The results are obtained by combining the simultaneous influence of two groups of factors (lifestyle and health care resources) on the efficiency of health systems. In the case of variables characterizing lifestyle factors, the criteria determining the assignment of individual factors as input or output variables are presented. In the discussion of the results, particular attention is given to the heterogeneity of the studied group of countries and the isotonic character of the set of variables. In this context, attention is drawn, among other things, to the so-called "efficiency by default." Moreover, using the projection results to verify the correctness of the obtained results is presented. The main practical implication of this study is that the conclusions based on the results can help decision-makers evaluate the activities of health systems and contribute to identifying directions for future improvement.

The rest of this article is organized as follows. The "Impact of lifestyle factors on health" section presents the relevance of non-medical factors on health systems. The "Methodology" section presents basic information about the non-parametric DEA method in the context of benchmarking and the NDEA algorithm used in the article. The "Model specification" section presents the structure of the variables and their interrelationships—in particular regarding the intermediate product, which plays a dual role in the model. The "Data" section contains information about the data sources, the period from which the data are derived, and basic descriptive statistics of individual variables. In the "Results" section, the preliminary results are presented in a concise form, and in the following "Discussion" section, the results are interpreted and explained. The article closes with the "Conclusion," where possibilities and directions of further research are indicated.

2. Impact of Lifestyle Factors on Health

Research on health systems usually focuses on determining the impact of health expenditure and medical infrastructure on public health, but the impact of non-medical factors—biological, socioeconomic, and lifestyle-related factors—should not be ignored (Hollingsworth, 2012; OECD, 2010; RWJF, 2019; Rettenmaier and Wang, 2013; González *et al.*, 2010). Biological factors include gender and age structure, particularly the proportion of the population over 65 years of age. Socioeconomic factors include a level of education, income, unemployment, the economic, social, and cultural status of the population, and environmental pollution resulting from the urbanization of the region of residence. Lifestyle factors such as smoking, alcohol consumption, dietary patterns leading to overweight and obesity, and physical inactivity significantly increase the risk of morbidity and mortality (Di Cesare *et al.*, 2013; Foster *et al.*, 2018; Cawley and Ruhm, 2012). According to the WHO (2019), the tobacco smoking epidemic is one of the most severe threats to public health, killing more than 8 million people per year, of which approximately 1.2 million deaths result from non-smokers' exposure to second-hand smoke. Di Cesare *et al.* (2013) and Foster *et al.* (2018) note that people with low socioeconomic status are at greater risk of health loss from non-communicable diseases than are those with higher socioeconomic rank.

Many health care outcomes do not result directly from systemic interventions but are influenced by the abovementioned non-medical factors (Retzlaff-Roberts *et al.*, 2004; Papanicolas and Cylus, 2017; OECD, 2010). Spinks and Hollingsworth (2009) have a similar opinion and state that the commonly used health outcome indicators based on life expectancy mainly reflect people's lifestyles and socioeconomic and environmental factors. An unhealthy lifestyle is associated with a higher risk of mortality. In contrast, the positive impact of a healthy lifestyle on life expectancy may increase the population's average age, which may increase the burden on health systems, affecting their efficiency (European Union, 2015).

Non-medical factors affecting health have been accounted for in different ways in previous studies. Retzlaff-Roberts *et al.* (2004) considered dietary choices, physical activity, and tobacco and alcohol use as exogenously fixed inputs since they were considered beyond the short-term discretionary control of policymakers. They used the DEA model with non-discretionary inputs; therefore, in the efficiency calculations, these variables remained unchanged. Another frequently used way of considering non-medical inputs is to conduct a two-stage analysis, (Afonso and Aubyn, 2011; Hadad *et al.*, 2013; de Cos and Moral-Benito, 2014), in which two different sets of variables are used. In the first stage, the set of variables underlying the health production function is used to estimate the efficiency indices according to the DEA method. In the second stage, the impact of non-medical factors, shaped and controlled by policymakers and influencing the functioning of health production processes, is taken into account by regressing efficiency scores on non-medical factors. The development of the NDEA method has created new opportunities to extend the analysis of health system efficiency to include non-medical factors.

Ozcan and Khushalani (2017), similarly to other authors, emphasized the importance of lifestyle-related factors on society's level of health. However, they differ in their treatment of lifestyle factors, stating that they are beyond the control of health care systems and are regulated and controlled by the government institutions responsible for public health in each country. This characteristic is reflected in legislation and actions related to health education, such as those concerning the harmfulness of smoking, alcohol consumption, and obesity. Ozcan and Khushalani (2017) proposed the NDEA model, creating a representation of a health system in the form of two divisions: public health and medical care. They included non-medical factors directly in the NDEA model as inputs to the public health division. The medical care division used a traditional resource-based approach, taking human resources and medical infrastructure as inputs. As an intermediate product combining these two divisions, they adopted preventive services such as vaccination and screening. The rationale for this approach is that properly conducted prevention activities, which are the responsibility of the public health system, reduce the burden on the health care system.

3. Methodology

As health becomes more important on the global agenda, there is a growing need to accurately measure its complex dimensions and assess the impact of changes in health policy. A good understanding of how health systems work enables policies to be appropriately framed and resources to be used in the best possible way. This can be achieved only if there is a firm foundation of metrics and evaluation methods (Hollingsworth, 2012).

Hollingsworth (2012) suggests that actions should be taken to measure the efficiency of health systems more applicable to recipients. Such analyses are required to produce valid and robust results, which can be achieved by taking into account relevant methodological requirements, including appropriate model specifications, consideration of sensitivity analysis and data testing in the construction of the model, and appropriate interpretation of the results, taking into account the importance of all key performance issues. This opinion is confirmed by Wendt (2014), who indicates that comparative studies are best for assessing the efficiency of similar health care systems in different countries. The above requirements are met by the many variants of the DEA method provided that an appropriate model is selected, the basic assumptions are met, and the DMUs are engaged in similar activities, produce comparable products or services (which enables us to define a standard set of outputs) using a similar range of resources, and operate in comparable environments (Avkiran, 2011; Dyson *et al.*, 2001; Courtis *et al.*, 2020).

According to Cook *et al.* (2014), when selecting a model, several key issues should be considered, such as the objective of the study, the DMUs to be compared, the inputs and outputs characterizing the DMUs, the isotonic nature of inputs and outputs returns to scale, the relationship between the number of DMUs to be compared and the total number of inputs and outputs, and the orientation of the model.

In its original version (Charnes *et al.*, 1978), DEA represented a production process in which the required resources are inputs, and the products are outputs. In such cases, the DEA model maps the processing of the inputs into outputs, and the result is a production frontier created by efficient DMUs. Despite the vital link between DEA and production theory in economics, the development of the method has led to its application in benchmarking. When DEA is used for benchmarking, specific characteristics of the objects to be compared that are relevant for the study are identified, as opposed to using the resources that are changed into products due to the transformation process. In such a case, the efficient DMUs defined by the DEA create the so-called frontier of best practices (Cook *et al.*, 2014). This issue is called the balanced benchmarking problem (Sherman and Zhu, 2013) or the general benchmarking problem (Cook *et al.*, 2014).

DEA is a multi-criteria methodology for evaluating alternative DMUs, and DEA inputs and outputs are two sets of performance criteria, where the set of inputs is to be minimized, and the set of outputs is maximized. In the case of the traditional approach based on the production process, the principle of isotonicity of inputs and outputs must be satisfied, i.e., increasing input values reduce efficiency, while increasing output values increase efficiency (Dyson *et al.*, 2001; Avkiran, 2011; Spinks and Hollingsworth, 2009; Mitropoulos, 2019). According to Bao *et al.* (2010), this condition is challenging to meet in exchange (trade) and many other areas. When applying DEA to benchmarking, the features describing DMUs do not represent resources and products in the standard production concept. In the benchmarking literature, terms such as indicators or measures are used. Thus, the problem is how to classify these measurements of performance into input and output categories for use in DEA (Cook *et al.*, 2014). If DEA is used for benchmarking, the inputs are performance measures of "the-fewer-the-better" type, whereas the outputs are performance measures of "the-more-the-better" type (Cook *et al.*, 2014; Afonso and Aubyn, 2011; Ouenniche *et al.*, 2014; Tone, 2017; Hadad *et al.*, 2013).

Another issue in formulating the DEA model is the existence of positive or negative returns to scale. A DMU may be too small to achieve optimal efficiency or too large, making it challenging to manage. If the variable returns scale (VRS) model is used when there are no inherent returns to scale, small and large DMUs will overestimate the efficiency assessment. The VRS model can be used only if returns to scale can be demonstrated (Dyson *et al.*, 2001). Ozcan and Khushalani (2017) conclude that the VRS model requires an a priori assumption about whether the examined health care systems have increasing or decreasing returns to scale. They could not make such assumptions due to the unavailability of confirmation in the literature; therefore, they used a constant return to scale (CRS) model. Country-by-country comparisons often use indicator variables to relate the values of the assessed factors to scaling variables such as GDP, population, and the number of employees (González *et al.*, 2010; Dosi *et al.*, 2006); therefore, the CRS model is justified.

A large number of inputs and outputs relative to the number of DMUs being compared reduces the discriminatory power of DEA. Golany and Roll (1998) suggest that the number of DMUs compared should be twice as large as the total number of inputs and outputs considered.

Depending on whether the inputs or outputs are controllable, the model's orientation towards the inputs or the outputs is assumed (Thanassoulis, 2001). This enables the assessment of the deficiencies of either the inputs or the outputs. Alternatively, a non-oriented model can also be used, for example, the non-oriented NDEA CRS model used by Ozcan and Khushalani (2017). Such models enable the assessment of deficiencies in the inputs, outputs, and links for inefficient DMUs.

The production process involves various interrelated activities, each of which has its exogenous input data and final results together with intermediate measures produced and consumed in the system. To take these factors into account, NDEA models are used in the DMUs under assessment to consider the impact of internal processes (Hatami-Marbini and Saati, 2019).

In this article, the slack-based network DEA model (NDEA-SBM) proposed by Tone and Tsutsui (2009) was used to account for input excesses and output shortfalls directly.

The subject of the efficiency analysis is *n* DMUs $(j = 1, \ldots, n)$ consisting of K divisions $(k = 1, \ldots, K)$. Let m_k and r_k be the numbers of inputs and outputs for division k, respectively. The link leading from division k to division h is denoted by (k,h), and a set of links is denoted by L. The data observed are $\{x_j^k \in R_+^{m_k}\}(j = 1, \ldots, n; k = 1, \ldots, K)$ (inputs to DMU_j in division k), $\{y_j^k \in R_+^{r_k}\}(j = 1, \ldots, n; k = 1, \ldots, K)$ (outputs from DMU_j in division k) and $\{z_j^{(k,h)} \in R_+^{t_{k,h}}\}(j = 1, \ldots, n; (k,h) \in L)$ (linking intermediate products from division k to division h), where $t_{(k,h)}$ is the number of items in the link (k,h).

$DMU_o\,(o = 1, \ldots, n)$ is represented by:

$$\begin{aligned} & x_o^k = X^k\lambda^k + s^{k-}(k = 1, \ldots, K) \\ & y_o^k = Y^k\lambda^k - s^{k+}(k = 1, \ldots, K) \\ & e\lambda^k = 1\ (k = 1, \ldots, K) \\ & \lambda^k \geq 0, \qquad s^{k-} \geq 0, \qquad s^{k+} \geq 0, \qquad (\forall k) \end{aligned} \tag{1}$$

where $\lambda^k \in R_+^n$ is the intensity vector corresponding to the division $k\ (k = 1, \ldots, K)$ and s^{k-} (s^{k+}) are vectors of input (output) slacks.

The linking activities are freely defined (discretionary), keeping continuity between the input and the output:

$$Z^{(k,h)}\lambda^{h} = Z^{(k,h)}\lambda^{k}, (\forall(k,h)) \tag{2}$$

The non-oriented efficiency $\mathrm{DMU_o}$ is calculated as follows:

$$\rho_o^* = \min_{\lambda^k, s^{k-}, s^{k+}} \frac{\sum_{k=1}^{K} w^k \left[1 - \frac{1}{m_k}\left(\sum_{i=1}^{m_k} \frac{s_i^{k-}}{x_{io}^k}\right)\right]}{\sum_{k=1}^{K} w^k \left[1 + \frac{1}{r_k}\left(\sum_{r=1}^{r_k} \frac{s_r^{k+}}{y_{ro}^k}\right)\right]} \tag{3}$$

where $\sum_{k=1}^{K} w^k = 1, w^k \geq 0 (\forall k)$ and w^k is the relative weight of division k, defined subjectively according to its significance.

A detailed description of the method used, including the method for calculating the efficiency of individual divisions, can be found in Tone and Tsutsui (2009) and publications describing the practical applications of NDEA, e.g., in the evaluation of the efficiency of technology development programs (Lu *et al.*, 2016).

4. Model Specification

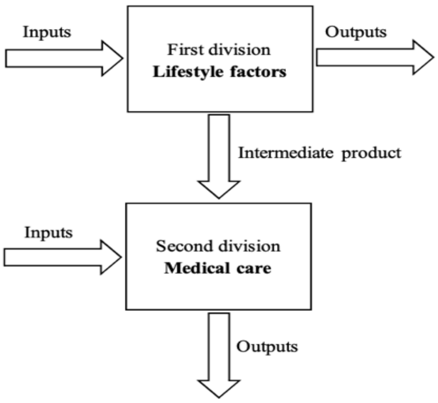
When formulating the empirical model, its specifications (Murillo-Zamorano and Petraglia, 2011), which must be consistent with the aim of the study, should be defined (Bao *et al.*, 2010). The first step is the selection of a model. This paper aims to assess the influence of two sets of factors on the health system's performance, so the NDEA-SBM model with CRS, described in the previous section, was chosen.

One of the main problems in building NDEA models is determining the intermediate product that links the divisions of the network model. Mitropoulos (2019) stresses that the dual nature of indirect dual-use measures leads to an inherent trade-off between the combined stages. This feature appears to represent a fundamental difficulty in the selection of intermediate variables in network models. When using a two-stage network structure, conflict may arise between two stages (Hatami-Marbini and Saati, 2019), whereby the second stage may require a reduction in intermediate measures (inputs) to achieve efficiency. However, such an action entails a reduction in the intermediate measures (outputs) of the first stage, thus reducing its efficiency. Ozcan and Khushalani (2017) used variables reflecting preventive measures, such as vaccination and screening for chronic disorders, in their health systems analysis. These variables act as outputs of the public health division and as inputs into the health care division. From the perspective of improving health, the values of these variables must be maximized rather than minimized. Thus, as outputs of the public health division, values that are as high as possible are desirable.

On the other hand, the same variables are inputs to the health care division, and their values should be minimized according to the DEA methodology. It is practically impossible for health systems to find variables that satisfy such input and output conditions. Ozcan and Khushalani (2017) explained the correctness of using these variables as an input into the health care division by explaining that the higher their values are, the lower the potential burden on the health care division. This is a compelling explanation.

Based on preliminary analyses of three possible model orientations (oriented towards inputs or outputs and non-oriented), a non-oriented model consisting of two divisions was selected. This model describes the investigated phenomenon in the most reliable manner. The first division reflects the influence of lifestyle factors on the health system's performance, and the second division reflects the influence of health system resources. A diagram of the model is shown in Figure 1.

Figure 1. *Model structure*



Source: *Own study.*

In the next step, the input, output, and intermediate product variables were selected from the broad set of data used in the initial tests. Alcohol consumption, tobacco smoking, and overweight were included as inputs to the first division: ALC_WEEK – a share of individuals claiming to drink alcohol every week; CURR_SMOKER – a share of individuals who are current smokers; OVERWEIGHT – a share of overweight individuals in the population. Self-perceived health of the population aged 65 years and over is included as an output in this division: S_VG_G_65 variable —share of individuals who assess their health to be very good or good. The selection of these variables was based on the analysis conducted in the "Impact of lifestyle factors on health" section. The inputs to the second division included primary health care resources, such as medical personnel, medical infrastructure, and financial outlays PHYS – number of physicians per hundred thousand inhabitants; NUR_MID – number of nurses and midwives per hundred thousand inhabitants; BEDS – number of hospital beds per hundred thousand inhabitants; EXP_TOT_GDP – total health care expenditure as a percentage of GDP. The outputs were the expected numbers of healthy life years in absolute value at 65 for women (HLE_65_F) and men

(HLE_65_M), as explained in the next section. The selection of variables was based on the analysis conducted in the introduction. Adopting a 65 years of age limit for outputs is motivated by the fact that this age group has a much higher need for medical care than do younger individuals.

The assignment of the abovementioned characteristics to the inputs or outputs was guided by the less-the-better principle for the inputs and the more-the-better principle for the outputs, as described in the previous section. One of the main difficulties in network models is finding a variable to serve as an intermediate product between divisions, i.e., a variable that acts as an output of the first division and an input of the second division. In this paper, EXP_PREV_GDP — expenditure on preventive care, expressed as a percentage of GDP — is used as the link between the two divisions. The choice of this variable was based on the argument given by Ozcan and Khushalani (2017) that correctly conducted prevention activities, which are the responsibility of the public health system, reduce the burden on the health care system. By analogy, more lavish spending on prevention should reduce the burden on the health care division. An additional justification for using this variable as a link is the use of a non-oriented model in these studies, which significantly minimizes the potential conflict resulting from the dual role of this variable. This characteristic was verified based on the projection of the variable being an intermediate product for the three different orientations of the model. A more detailed description can be found in the “Discussion” section.

Except for HLE_65_F and HLE_65_M, the variables used in the model are indicators, so the final results are not affected by the returns to scale due to the size of the country (this problem is reported in the “Methodology” section (Dyson *et al.,* 2001; González *et al.,* 2010).

Cook *et al.* (2014) state that the DEA score should not be referred to as “production efficiency within the framework of general benchmarking.” In such cases, it should the DEA score represents the “overall performance” of the unit. However, the model presented in this paper is mainly benchmarking in nature in the lifestyle factor division but productive in the medical care division, where the inputs are resources transformed into health outcomes. Therefore, in the following part of this paper, the term “efficiency score” will be used to evaluate both the individual divisions and the entire health system.

5. Data

The data used in this paper are primarily derived from the Eurostat database (EUROSTAT, 2019), which provides administrative health data on an annual basis. The self-assessment data (e.g., self-assessment of health and lifestyle factors) from the European Health Interview Survey (EHIS) and the European Union Statistics on Income and Living Conditions (EU-SILC), which are conducted every five years, are also published. Therefore, the administrative data are from 2016, and the survey data are from 2014. The survey covered 30 countries — all EU28 countries supplemented by Iceland and Norway. The last year for which the data were most complete was selected. A small number of missing data (23 of 341 items used in the calculations) were supplemented with values from the closest year. In one case, the value of the ALC_WEEK variable for the Netherlands, due to lack of any data, was imputed as the EU mean value via unconditional mean imputation (OECD 2008). Imputation of missing data is a common technique; however, according to Ozcan and Khushalani (2017), the approach may introduce a minor amount of bias in the findings. Regardless, entering only one value for the whole dataset does not substantially impact, as confirmed by the results of a study called "Death due to alcoholic abuse" published by Eurostat (EUROSTAT 2019a). The share of deaths due to alcohol abuse in the Netherlands is at the mean for EU countries, and the imputation enables the Netherlands to remain in the set of countries to be compared. Descriptive statistics for all variables are presented in Table 1.

Table 1. *Descriptive statistics of the variables included in the analyses*

Variable	Role	Year	Division	Mean	SD	Min	Max
ALC_WEEK	Input	2014	1	26.0	9.9	12.9	45.0
CURR_SMOKER	Input	2014	1	24.8	4.5	16.7	34.8
OVERWEIGHT	Input	2014	1	51.6	3.9	43.8	59.6
S_VG_G_65	Output	2016	1	36.4	17.6	5.7	65.8
PHYS	Input	2016	2	355.3	65.0	241.6	513.0
NUR_MID	Input	2016	2	923.6	377.0	350.0	1,804.9
BEDS	Input	2016	2	486.5	168.6	233.9	806.3
EXP_TOT_GDP	Input	2016	2	8.5	1.7	5.0	11.5
HLE_65_F	Output	2016	2	9.5	3.2	4.2	16.6
HLE_65_M	Output	2016	2	9.4	3.0	4.4	15.5
EXP_PREV_GDP	Link	2016	1-2	0.22	0.10	0.08	0.52

Source: *Own calculations.*

The lifestyle factors used vary widely from country to country. One in four surveyed countries is currently a smoker, and 26% drink alcohol every week. The proportion of smokers ranges from 16.7% in Sweden to 34.8% in Bulgaria, and the proportion of alcohol drinkers ranges from 12.9% in Lithuania to 45% in the UK. Therefore, habits vary widely, and prevention policies have very different efficiencies. The problem of obesity appears to be much more severe. On average, every second resident is overweight, and the proportion is similar in all the countries surveyed (from 43.8% in Italy to 59.6% in Malta). Thus, the obesity epidemic is independent of the country. However, the most significant differences are observed in the self-assessment of health: on average, 36.4% of all respondents from all countries assess their health as very good or good. In Lithuania, this opinion is expressed by only 5.7% of respondents, whereas 65.8% of respondents in Ireland express this opinion. Such variation may be attributed to the significant impact of lifestyle factors and healthcare systems' potential weaknesses.

The set of variables describing the functioning of health care also includes essential differences between countries. The maximum to minimum values of inputs representing human resources and infrastructure are as follows: PHYS, 2.1; NUR_MID, 5.2; and BEDS, 3.4. A significant factor is health care spending, which averaged 8.5% of GDP overall, 4.4% in Latvia, and 15.5% in Iceland. For outputs, two variables, HLE_65_F, and HLE_65_M were adopted due to significant differences between women and men in different countries in terms of the frequently used life expectancy at birth (LE), which is 5.9 years longer for women in all EU countries on average (for 2016), with the minimum value of 3.2 years in the Netherlands and the maximum value of 10.2 years in Lithuania. HLE_65 is 0.0–1.5 years higher for women in half of the countries surveyed and 0.2–1.5 years higher for men in the remaining half of the countries. Therefore, the adoption of values without a gender division is not justified.

The EXP_PREV_GDP variable, which links the divisions, indirectly maps the activity of prevention policies, including measures to minimize the impact of harmful lifestyle factors. The average value is 0.22%, but significant variation is observed across countries, from 0.08% in Romania to 0.52% in the UK.

6. Results

Table 2. *Efficiency scores for European countries*

Country/ Abbreviation		OV		LF			MC		
		Score	Rank	Score	Rank	Peers	Score	Rank	Peers
Austria	AT	0.509	20	0.601	14		0.450	26	
Belgium	BE	0.695	13	0.715	9		0.676	14	
Bulgaria	BG	0.471	21	0.290	26		1	1	0
Croatia	HR	0.336	26	0.302	24		0.414	28	
Cyprus	CY	1	1	1	1	2	1	1	2
Czech Republic	CZ	0.390	23	0.299	25		0.662	16	
Denmark	DK	0.747	10	0.796	8		0.702	13	
Estonia	EE	0.337	25	0.273	27		0.527	24	
Finland	FI	0.607	16	0.663	11		0.558	21	
France	FR	0.670	14	0.694	10		0.643	18	
Germany	DE	0.567	17	0.541	17		0.602	19	
Greece	EL	0.800	8	0.600	15		1	1	0
Hungary	HU	0.371	24	0.311	23		0.516	25	
Iceland	IS	1	1	1	1	18	1	1	3
Ireland	IE	1	1	1	1	0	1	1	0
Italy	IT	0.728	12	0.588	16		0.949	11	
Latvia	LV	0.218	29	0.166	29		0.435	27	
Lithuania	LT	0.147	30	0.109	30		0.393	29	
Luxembourg	LU	0.643	15	0.639	12		0.647	17	
Malta	MT	0.741	11	0.481	18		1	1	0
Netherlands	NL	0.813	7	1	1	2	0.669	15	
Norway	NO	0.898	6	1	1	10	0.801	12	
Poland	PL	0.522	18	0.320	22		1	1	0
Portugal	PT	0.282	28	0.211	28		0.558	22	

Romania	RO	0.436	22	0.380	20		0.541	23	
Slovakia	SK	0.333	27	0.329	21		0.340	30	
Slovenia	SI	0.511	19	0.468	19		0.581	20	
Spain	ES	0.791	9	0.629	13		1	1	1
Sweden	SE	1	1	1	1	0	1	1	20
United Kingdom	UK	1	1	1	1	0	1	1	10
Mean		0.619		0.580			0.722		
SD		0.256		0.291			0.232		
Min		0.147		0.109			0.340		

Source: *Own calculations.*

The efficiency scores for health systems (OV), calculated based on the NDEA-SBM model with CRS, described by equation (3), and efficiency results for two divisions, i.e., lifestyle factors (LF) and medical care (MC), are shown in the columns headed "Score" in Table 2. For these calculations, the same weight value of 0.5 was used in equation (3) for both divisions. The values of the efficiency scores are complemented by the ranking positions presented in the "Rank" column. For 100% efficient countries, the "Peers" column shows how many times they were used as a benchmark for inefficient countries. The calculations were performed using MaxDEA Ultra 6.19 software.

The health systems are fully efficient in five countries (CY, IS, IE, SE, and the UK). These countries also achieve total efficiency for both LF and MC divisions. The average efficiency of the health systems in all countries is relatively low at 0.619: Lithuania has the lowest value of 0.147, mainly due to the low efficiency of the LF division of only 0.109. The average efficiency for this division is 0.580: for 13 countries, it is less than 0.5, and for six countries, it is even less than 0.3. In addition to the five countries with fully efficient health systems, two countries (NL and NO) are fully efficient in this division. In the MC division, the average efficiency for all countries is 0.72, which is not high, but the variation between countries is more minor. Slovakia has a minimum value of 0.340. Total efficiency was achieved in 10 countries (BG, CY, EL, IS, IE, MT, PL, ES, SE, UK).

7. Discussion

The relatively low average efficiency is partly due to the structure of the compared group of countries. The seventeen old EU countries, as well as Iceland and Norway, are countries with a higher level of economic development than that of the eleven new countries (BG, HR, CZ, EE, HU, LV, LT, PL, RO, SK, and SI), which undoubtedly affects the efficiency of health systems. This feature is illustrated in Table 3, where the "NEW" column shows the statistics calculated for the 11 new EU countries, and the "OLD" column presents the statistics for the remaining countries. These results are confirmed by Foster *et al.* (2018), who found that populations with lower socioeconomic status have worse health outcomes than those observed in affluent populations.

Table 3. *Descriptive statistics of the efficiency scores (OLD and NEW countries)*

	OLD			NEW		
	OV	LF	MC	OV	LF	MC
Mean	0.763	0.745	0.803	0.370	0.295	0.583
SD	0.196	0.230	0.199	0.116	0.095	0.226
Min	0.282	0.211	0.450	0.147	0.109	0.340

Source: *Own calculations.*

The average efficiency score (OV) of the health systems in the new EU countries is twice as low as that in the remaining ones, mainly due to the very low efficiency of the LF division, the average efficiency of which for new countries is 2.5 times lower than that of the old EU counterparts. The situation is slightly better in the MC division, where the average efficiency of the new EU countries is only 1.4 times lower. The minimum efficiency values are similar mainly because of the values of all the direct outputs and the intermediate product and the values of some inputs that are less favorable for the new EU countries. The descriptive statistics for these variables for the new and old countries are presented in Table 4.

Table 4. *Descriptive statistics of the variables included in the analyses (OLD and NEW countries)*

Statistics	ALC_WEEK	CURR_SMOKER	OVERWEIGHT	S_VG_G_65	PHYS	NUR_MID	BEDS	EXP_TOT_GDP	HLE_65_F	HLE_65_M	EXP_PREV_GDP
	OLD										
Mean	30.3	23.0	50.2	46.8	370.7	1050.9	412.7	9.4	11.0	11.0	0.252
SD	9.2	4.4	4.4	13.3	64.2	418.7	163.5	1.5	2.8	2.3	0.112
Min	16.7	16.7	43.8	12.1	278.3	350.0	233.9	6.2	6.4	7.7	0.080
Max	45.0	32.6	59.6	65.8	513.0	1804.9	806.3	11.5	16.6	15.5	0.520
	NEW										
Mean	18.5	27.9	54.0	18.4	328.6	703.8	614.0	7.0	6.8	6.6	0.18
SD	6.5	2.9	1.2	6.9	63.4	163.9	94.5	0.9	2.0	1.7	0.06
Min	12.9	24.2	52.3	5.7	241.6	480.2	448.7	5.0	4.2	4.4	0.08
Max	34.6	34.8	55.8	32.2	446.7	974.0	727.0	8.5	10.1	9.2	0.25

Source: *Own calculations.*

The direct output from the LF division is the subjective health assessment S_VG_G_65. The average and maximum for new EU countries are 2.5 times and twice as low as those of the remaining EU members. Similarly, the expected number of healthy life years at 65 for women and men (direct outputs from the MC division), HLE_65_F and HLE_65_M, are 1.6 times lower for the new countries. These factors are the primary outcomes of the health system in this model. One direct reason for this result is the lower health expenditure EXP_TOT_GDP in new EU countries, which is, on average, 26% of that in the remaining countries.

Similarly, expenditure on preventive care EXP_PREV_GDP (an intermediate product between the LF and MC divisions) is approximately 30% lower in new EU countries. The difference is not as evident in the case of the inputs to the MC division. The average number of doctors, PHYS, is comparable in old and new countries; the average number of beds, BEDS, in new countries is nearly 50% higher than that in other countries; and the average number of nurses and midwives, NUR_MID, is 33% lower in the new countries. Therefore, the structure, organization, and functioning of health care systems in new EU countries differ significantly from those in the remaining countries. The resources used in health care are probably superfluous, which is not conducive to their optimal use and may cause inadequate efficiency. This topic will be presented more extensively when discussing the projection calculated for inefficient countries.

In terms of the lifestyle factors included in the LF division, the alcohol and tobacco consumption variables characterize the habits of society but do not reflect the amounts consumed. The average proportion of smokers, CURR_SMOKER, is approximately 21% higher in new EU countries than in the remaining countries. In the case of alcohol consumption, ALC_WEEK, in new EU countries, the proportion is 39% lower than in the remaining countries. However, for the OVERWEIGHT variable, the data for both groups of countries are comparable.

Some of the above comparisons explain the differences in efficiency directly; however, some inputs, such as BEDS and ALC_WEEK, are not obvious. Thus, it is not possible to conclude the efficiency of a complex system based on individual indicators. Only the aggregation of all factors via DEA allows for the complete picture of the phenomenon to be obtained.

Nevertheless, the results of many factors aggregated via DEA must be interpreted with caution. Hadad *et al.* (2013); Spinks and Hollingsworth (2009); de Cos and Moral-Benito (2014) due to certain limitations of this non-parametric method. The location and shape of the efficiency limit in DEA are determined empirically, although this limit is sensitive to DMUs with unusual types, levels, or combinations of inputs or outputs (Hollingsworth and Street, 2006). Thus, one of the main pitfalls in DEA applications is placing nonhomogeneous DMUs in the same sample, distorting the results (Avkiran, 2011; Dyson *et al.*, 2001; Puig-Junoy, 1998; European Union, 2015). One way to ensure homogeneity of the sample is to exclude outliers (Puig-Junoy, 1998; Afonso and Aubyn, 2005; Hadad *et al.*, 2013), which prevents a complete comparison of the tested group of units. Another approach is clustering units into homogeneous sets and running separate DEAs (Avkiran, 2011; Dyson *et al.*, 2001), which reduces the size of the tested samples and thus limits the number of usable variables and fails to provide a complete comparison of the whole group. DEA uses a selective amount of data to estimate the efficiency outcome of each DMU by comparing each DMU only with peers that produce a comparable output mix. If outlier observations exist, some DMUs may not have any peers, which results in the automatic assignment of total efficiency to the considered DMU (Hollingsworth and Street, 2006). Such DMUs are considered efficient by default, which means that they are not dominated by other DMUs but do not dominate any other DMUs (Afonso and Aubyn, 2005).

However, leaving such units in the sample does not affect the outcomes of other DMUs (Afonso and Aubyn, 2005). Hence, in some studies, nonhomogeneous DMUs were included in the sample, which requires additional careful interpretation of the results. In surveys of health care systems of OECD countries, Retzlaff-Roberts *et al.* (2004) found that both countries with good health outcomes and countries with modest or relatively poor health outcomes, which in their opinion, may be a model for more economical allocation of health care resources, are efficient. In the interpretation of the technical efficiency outcomes of health systems in EU countries (European Union, 2015), attention is drawn to the excellent efficiency outcomes of Romania, which are difficult to explain. The small number of peers can explain this result and hence is an outlier that has been artificially placed at the frontier.

Similar results were obtained by Mitropoulos (2019) in the production division of the NDEA model, where countries with relatively poor health outcomes, such as Romania and Bulgaria, achieved total efficiency.

The “Peers” column of Table 2 includes three countries “efficient by default” in the LF division (IE, SE, and the UK) and five such countries in the MC division (BG, EL, IE, MT, and PL). The values of the variables for these countries and descriptive statistics for all countries are presented in Table 5. An unusual set of values of inputs and outputs is observed in the LF division. These countries have very high values of the S_VG_G_65 direct output (IE has the maximum value, and SE and UK have values close to the maximum). The same is true for the EXP_PREV_GDP intermediate product (the UK has the maximum value and SE and IE have values close to the maximum). CURR_SMOKER has the minimum value for the UK on the input side, and the values for SE and IE are close to the minimum. ALC_WEEK has the maximum value in the UK, and SE and IE have values close to the maximum. By contrast, OVERWEIGHT does not influence the efficiency outcome.

Table 5. *Variable values for the “efficient by default” countries*

Country	ALC_WEEK	CURR_SMOKER	OVERWEIGHT	S_VG_G_65	PHYS	NUR_MID	BEDS	EXP_TOT_GDP	HLE_65_F	HLE_65_M	EXP_PREV_GDP
	Lifestyle division				Medical care division						Link
BG					413.8	480.2	137.6	8.2	10.1	9.2	0.21
EL					459.2	350.0	238.0	8.5	7.8	8.0	0.11

MT					382.8	860.0	214.1	9.3	12.9	12.8	0.12
PL					241.6	578.8	150.6	6.5	8.9	8.2	0.20
IE	41.4	22.0	54.4	65.8	293.5	1407.6	384.8	7.4	13.2	12.0	0.24
SE	36.9	16.7	47.7	61.0							0.34
UK	45.0	17.3	55.0	53.8							0.52
Mean	26.0	24.8	51.6	36.4	355.3	923.6	234.5	8.5	9.5	9.4	0.22
Min	12.9	16.7	43.8	5.7	241.6	350.0	124.0	5.0	4.2	4.4	0.08
Max	45.0	34.8	59.6	65.8	513.0	1804.9	427.6	11.5	16.6	15.5	0.52

Source: *Own calculations.*

An unusual set of values of inputs and outputs is also observed in the MC division. Three countries, BG, IE, and MT have HLE_65_F and HLE_65_M values that are significantly higher than the average while having an average EXP_TOT_GDP. A significant variation in the values of the remaining inputs is observed. IE has a high NUR_MID close to the maximum, whereas that of MT is close to average, and that of BG is equal to half of the average. For PHYS and BEDS, this variation is more minor, and the countries rank, in descending order, as follows: BG, MT, and IE. For BEDS, the ranking is IE, MT, and BG. Such differentiation in inputs with comparable outputs is the main reason these countries are not directly comparable. The two remaining countries, EL and PL, are countries with relatively weak outputs and invalid inputs. Significant differences can also be identified here because these countries are not comparable. The value of PHYS for EL is almost twice as high as that for PL, the value of BEDS is 1.7 times higher for EL than for PL, and the value of NUR_MID is 1.7 times higher for PL than for EL. Substantial differences are also observed for EXP_PREV_GDP, where for EL and MT, the values are nearly twice as low as those for BG, PL, and IE.

None of the "efficient by default" countries can serve as a model for formulating recommendations for other inefficient countries, considered when calculating the projections. An integral part of DEA outcomes is the projection of variable values for inefficient countries based on the best practices of their fully efficient counterparts. Projections can be used to determine whether inefficient countries should modify their inputs or outputs (depending on the model's orientation) to achieve total efficiency. A projection can also help to choose the orientation of the model for final calculations. This article calculates the efficiency results for input-, output- and non-oriented models. Regardless of the model's orientation, the same countries achieve total efficiency, both overall and in individual divisions. The level of inefficiency in the remaining countries varies. The extent to which the model orientation influences the projected value of the intermediate product EXP_PREV_GDP was assessed in the first step of the projection analysis. The average value of this variable based on the data is 0.225, while in the input-oriented model, the average projection is 0.177, in the output-oriented model, the projection is 0.236, and in the non-oriented model, the projection is 0.220. For example, in the input-oriented model for IT, the variable EXP_PREV_GDP is reduced from 0.360 to 0.127. Such a significant reduction in prevention expenditure appears to be unjustified. These outcomes lead to choosing a non-oriented model that relates to the value of this variable in the most neutral way.

The second problem that can be assessed based on the projection results is the scale of output changes, which in the case of the data set used in this study are directly non-controllable. This characteristic can be illustrated by the example of the HLE_65_F and HLE_65_M variables, which have average values of 9.5 and 9.4 in the recorded data and 11.8 and 11.4 in the projection for the output-oriented model. By contrast, in the non-oriented model, the respective values are 10.5 and 9.9. These examples illustrate the risk of obtaining difficult results to explain and use if the model is not oriented correctly. The projections appear to be most stable and reliable for the non-oriented model; thus, such a model has been chosen.

The last step is to check whether the direction of change in the projections follows the isotonic principle. Since the projection brings all DMUs to total efficiency, the outputs should be increased in the projections, and the inputs should be decreased, whereas, in the case of the intermediate products, changes should be made in both directions due to the dual nature of this variable. For the non-oriented model in the projection for inputs, no increase cases were observed; for outputs, no decrease cases were observed. By contrast, for the intermediate product, a decrease in value was observed for 11 countries, an increase was observed for seven countries, and no change was observed for seven countries. These results confirm that the choice of the uninformed model is correct.

The projection results may suggest directions to shape public health policy for inefficient countries, especially in terms of LF. Table 6 presents a summary of the values of the variables recorded in the five least-efficient countries (rows A) and the levels of desired values (rows P) that would allow health systems to operate efficiently at the level of the leaders of the previously presented ranking.

Table 6. *Projection for selected countries*

Country	Actual/ Projection	ALC_WEEK	CURR_SMOKER	OVERWEIGHT	S_VG_G_65	PHYS	NUR_MID	BEDS	EXP_TOT_GDP	HLE_65_F	HLE_65_M	EXP_PREV_GDP
		Lifestyle division				Medical care division						Link
LT	A	12.9	25.0	53.2	5.7	446.7	802.6	669.2	6.7	5.6	5.6	0.14
	P	12.9	12.4	36.6	38.2	158.0	438.4	88.8	4.1	6.1	5.6	0.13
LV	A	13.4	29.5	55.2	9.3	321.3	484.3	572.0	6.2	4.5	4.4	0.15
	P	13.4	12.9	38.0	39.7	122.3	347.1	81.1	3.5	4.8	4.4	0.14
PT	A	19.1	20.0	52.2	12.1	333.6	637.3	342.2	9.1	6.4	7.7	0.16
	P	18.4	17.7	52.2	54.6	216.9	603.3	124.5	5.7	8.4	7.7	0.19
SK	A	15.9	29.5	53.0	19.6	246.6	917.3	578.4	7.1	4.2	4.5	0.08
	P	15.9	15.3	45.1	47.1	204.8	566.3	112.2	5.2	8.0	7.2	0.16
HR	A	14.9	28.7	55.8	17.7	323.7	673.4	549.3	7.2	4.9	5.2	0.21

	P	14.9	14.4	42.3	44.2	145.1	409.4	92.6	4.0	5.7	5.2	0.15

Source: *Own calculations.*

As the non-oriented model allows for assessing the inefficiency of both inputs and outputs (and also of the intermediate product in network models), the projection concerns all the variables in the model. The reference set for these inefficient countries in the lifestyle section is, in all cases, IS, whereas in the health care section, for SK, it is SE, and for the remaining countries, SE and UK. Thus, the countries included in the reference sets determine the projection values. For the inputs of the lifestyle division, the most significant corrections are made to the CURR_SMOKER variable, where half should reduce the number of active smokers in four cases. More minor changes are suggested for OVERWEIGHT, while for ALC_WEEK, only for one country, PT, is a slight change suggested. Introducing such changes should improve the self-assessment of health S_VG_G_65 to a radical degree (e.g., for LT, an increase of almost a factor of seven). Changing the habits of the societies of these countries is a long-term process. For the inputs of the MC division, the projection shows that the health care system in the new countries involves far too many resources, including expenditure, about the outputs achieved. This issue is mainly due to differences in how health care systems are organized in the old and new countries, which is particularly visible for the BEDS variable. Such a radical change results from the fact that the projection is based on ES and UK practices, where the values of the BEDS variable are 234 and 258, respectively, with HLE_65_F outputs equal to 16.6 11.1 and HLE_65_M equal to 15.1 and 10.4, respectively. In the five most inefficient countries, HLE_65_F is between 4.2 and 6.4, and HLE_65_M is between 4.4 and 7.7. In this case, the optimization is conducted to minimize inputs while keeping the output values practically unchanged. As mentioned above, units that generate weak outputs at low cost can also achieve total efficiency.

The projection obtained from the model must be interpreted critically. For LT, LV, and HR, it should not be concluded that the BEDS variable should be limited to the level of 80–90. The results provide a message to policymakers that the resources involved in health care in these countries are disproportionate to the results. The aim is not to reduce the number of beds considerably but to introduce changes in the structure of the health care systems in these countries, following the best practices of SE and the UK.

8. Conclusion

The comparative analysis of health systems in different countries provides a better understanding of their performance, which is essential from undertaking activities related to their improvement. Improvement measures are necessary because health systems play a vital role in each country, influencing the level of safety perceived by society and the quality of life in general. The fact that health expenditure is one of the main areas of public spending is not insignificant. The complexity of these systems makes it challenging to define them precisely, which means there is no uniform template for their analysis.

The population health outcomes depend not only on the efficiency of the health care system but also on the level of commitment of resources, such as hospitals, doctors, and other medical infrastructure. Past and present lifestyle behaviors and environmental factors are also crucial. The aging population is not without significance in terms of the burden on the health care system. As a result, health system analyses are multidimensional and require the use of appropriate methods. In this article, the NDEA model was used, thereby enabling two groups of factors. The network structure of the model consists of two divisions, namely, LF and MC, which are connected by an intermediate product, i.e., the prevention expenditure. Such a structure allows the model to assess the efficiency of both the entire complex system and its divisions.

In all new EU countries, the lifestyle pulls their total efficiency down, while only a few the old EU takes place. Therefore, long-term actions promoting a healthy lifestyle, incredibly limiting tobacco smoking, are necessary. Improving lifestyle is costly and over time, and it does not bring direct financial benefits in reducing healthcare spending. As the statistics show, wealthier societies are usually also healthier.

The second variable (output) that requires radical improvement is the self-esteem of health. This variable also reduces the technical efficiency of the new EU countries. The reasons for very low self-esteem of health status among citizens of these countries may be a large percentage of people who feel no sense of economic security, low purchasing power of earnings, unhealthy diet, low physical activity, poor organization of health care - especially difficulties with quick medical appointments.

The new EU countries' problem is the relatively low number of medical personnel with a relatively extensive infrastructure (the BEDS variable). Relatively high expenditures with common health effects result in a low assessment of the health systems' effectiveness.

In this article, only LF and, indirectly, population aging is considered among the non-medical factors influencing health outcomes. Further research should examine other factors affecting the health system, such as environmental pollution and the socio-economic situation of the population, including unemployment and poverty levels. The level of unmet medical needs and the cause of this phenomenon is also an important consideration.

References:

Afonso, A., Aubyn, M. 2011. Assessing health efficiency across countries with a two-step and bootstrap analysis. Applied Economics Letters, 18(15), 1427-1430. doi: 10.1080/13504851.2010.541149.

Afonso, A, Aubyn, M. 2005. Non-parametric approaches to education and health efficiency in OECD countries. Journal of Applied Economics, 8(2), 227-246. doi.org/10.1080/15140326.2005.12040626.

Avkiran, N. 2011. Applications of data envelopment analysis in the service sector. In: Cooper, W., Seiford, L., Zhu, J. ed., Handbook on data envelopment analysis. (Springer, New York), 403-443.

Bao, C.P., Lee, K.C., Pu. H.L. 2010. A study of the input/output variable characteristics in DEA. Journal of the Chinese Institute of Industrial Engineers, 27(6), 429-437. doi:10.1080/10170669.2010.525393.

Cantor, V., Poh, K. 2018. Integrated analysis of healthcare efficiency: A systematic review. Journal of Medical Systems, 42(1). doi:10.1007/s10916-017-0848-7.

Cawley, J., Ruhm, C. 2011. The economics of risky health behaviors. in Pauly, M., McGuire, T., Barros, P. ed., Handbook of Health Economics. Vol. 2. (Elsevier, New York) 95-199.

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

Cheng, G., Zervopoulos, P. 2014. Estimating the technical efficiency of health care systems: a cross-country comparison using the directional distance function. European Journal of Operational Research, 238(3), 899-910. doi:10.1016/j.ejor.2014.05.007.

Cook, W., Tone, K., Zhu, J. 2014. Data envelopment analysis: Prior to choosing a model. Omega, 44, 1-4. doi:10.1016/j.omega.2013.09.004.

Curtis, P., Hanias, M., Kourtis, E., Kourtis, M. 2020. Data Envelopment Analysis (DEA) and Financial Ratios: A Pro-Stakeholders' View of Performance Measurement for Sustainable Value Creation of the Wind Energy. International Journal of Economics and Business Administration, 8(2), 326-350. DOI: 10.35808/ijeba/465.

de Cos, P., Moral-Benito, E. 2014. Determinants of health-system efficiency: evidence from OECD countries. International Journal of Health Care Finance and Economics, 14(1), 69-93. doi:10.1007/s10754-013-9140-7.

Di Cesare et al. 2013. Inequalities in non-communicable diseases and effective responses. Lancet, 381(9866), 585-597. doi:10.1016/S0140-6736(12)61851-0.

Dosi, G., Llerena, P., Labini, M. 2006. The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called 'European Paradox'. Research. Policy, 35(10), 1450-164. doi:10.1016/j.respol.2006.09.012.

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

European Union. 2015. Comparative efficiency of health systems, corrected for selected lifestyle factors. Final report. Retrieved from: https://ec.europa.eu/health/sites/health/files/systems_performance_assessment/docs/2015_maceli_report_en.pdf.

EUROSTAT. 2019. Population and social conditions/ Health. Retrieved from: https://ec.europa.eu/eurostat/data/database.

EUROSTAT. 2019a. Death due to alcoholic abuse. Retrieved from: https://ec.europa.eu/eurostat/databrowser/view/tps00140/default/table?lang=en.

Färe, R., Grosskopf, S. 2000. Network DEA. Socio-Economic Planning Sciencies, 34(1), 35-49. doi:10.1016/S0038-0121(99)00012-9.

Foster, H. et.al. 2018.The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a

prospective analysis of the UK Biobank cohort. The Lancet Public Health, 3(12), e576-e585. doi:10.1016/S2468-2667(18)30200-7.
Golany, B., Roll, Y. 1989. An application procedure for DEA. Omega, 17(3), 237-250. doi:10.1016/0305-0483(89)90029-7.
González, E., Cárcaba, A., Ventura, J. 2010. Value efficiency analysis of health systems: does public financing play a role? Journal of Public Health, 18(4), 337-350. doi:10.1007/s10389-009-0311-4.
Hadad, S., Hadad, Y., Simon-Tuval, T. 2013. Determinants of healthcare system's efficiency in OECD countries. The European Journal of Health Economics, 14(2), 253-265. doi:10.1007/s10198-011-0366-3.
Hatami-Marbini, A., Saati, S. 2020. Measuring performance with common weights:network DEA. Neural Computing and Applications, 32, 3599-3617. doi:10.1007/s00521-019-04219-4.
Hollingsworth, B. 2012. Revolution, evolution, or status quo? Guidelines for efficiencymeasurement in health care. Journal of Productivity Analysis, 37(1), 1-5. doi:10.1007/s11123-011-0221-7.
Hollingsworth, B., Street, A. 2006. The market for efficiency analysis of health care organizations. Health Economics, 15(10), 1055-1059. doi.10.1002/hec.1169.
Kao, C. 2014. Network data envelopment analysis: A review. European Journal of Operational Research, 239(1), 1-16. doi.10.1016/j.ejor.2014.02.039.
Kohl, S., Schoenfelder, J., Fügener, A., Brunner, J. 2019. The use of data envelopment analysis (DEA) in healthcare with a focus on hospitals. Health Care Management Science, 22(2), 245-286. doi:10.1007/s10729-018-9436-8.
Lu, W., Kweh, Q., Nourani, M., Huang, F. 2016. Evaluating the efficiency of dual-usetechnology development programs from the R&D and socio-economic perspectives. Omega, 62, 82-92. doi:10.1016/j.omega.2015.08.011.
Medeiros, J., Schwierz, C. 2015. Efficiency estimates of health care systems in the EU. European Commission, Economic Papers 549.
Mitropoulos, P. 2019. Production and quality performance of healthcare services in EU countries during the economic crisis. Operational Research. doi:10.1007/s12351-019-00483-3.
Mocan, N., Altindag, D. 2014. Education, cognition, health knowledge, and health behavior.The European Journal of Health Economics, 15(3), 265-279. doi:10.1007/s10198-013-0473-4.
Murillo-Zamorano, L., Petraglia, C. 2011. Technical efficiency in primary health care:does quality matter? The European Journal of Health Economics, 12(2), 115-125. doi:10.1007/s10198-010-0232-8.
OECD. 2010. Health care systems: Efficiency and policy settings. OECD Publishing, Paris.
OECD. 2008. Handbook on constructing composite indicators. Methodology and user guide. OECD Publications, Paris.
Ouenniche, J., Xu, B., Tone, K. 2014. Forecasting models evaluation using a slacks-basedcontext-dependent DEA framework. Journal of Applied Business Research, 30(5), 1477-1484. doi.org/10.19030/jabr.v30i5.8800.
Ozcan, Y., Khushalani, J. 2017. Assessing efficiency of public health and medical care provision in OECD countries after a decade of reform. Central European Journal of Operations Research, 25(2), 325-43. doi.org/10.1007/s10100-016-0440-0.
Papanicolas, I., Cylus, J. 2017. The challenges of using cross-national comparisonsof efficiency to inform health policy. Eurohealth, 23(2), 8-11.

Papanicolas, I., Smith, P. (ed.). 2013. Health system performance comparison. An agendafor policy, information and research. McGraw-Hill Education, Open University Press, New York.
Puig-Junoy, J. 1998. Measuring health production performance in the OECD. Applied Economics Letters, 5(4), 255-259. doi:10.1080/135048598354933.
Rettenmaier, A., Wang, Z. 2013. What determines health: a causal analysis using countylevel data. The European Journal of Health Economics, 14(5), 821-834. doi.10.1007/s10198-012-0429-0.
Retzlaff-Roberts, D., Chang, C., Rubin, R. 2004. Technical efficiency in the use of healthcare resources: a comparison of OECD countries. Health Policy, 69(1), 55-72. doi:10.1016/j.healthpol.2003.12.002.
RWJF. 2019. County health rankings. Key findings report. Retrieved from: https://www.countyhealthrankings.org/what-is-health.
Sherman, H., Zhu, J. 2013. Analyzing performance in service organizations. MIT SloanManagement Review, 54(4), 37-42.
Spinks, J., Hollingsworth, B. 2009. Cross-country comparisons of technical efficiencyof health production: a demonstration of pitfalls. Applied. Economics, 41(4), 417-427. doi:10.1080/00036840701604354.
Thanassoulis, E. 2001. Introduction to the theory and application of data envelopment analysis. A foundation text with integrated software. Springer New York.
Tone, K. (ed.). 2017. Advances in DEA theory and applications: With extensions to forecasting models.: John Wiley & Sons Ltd, New York. doi:10.1002/9781118946688.
Tone, K., Tsutsui, M. 2009. Network DEA: A slacks-based measure approach. European Journal Operational Research, 197(1), 243-52. doi.10.1016/j.ejor.2008.05.027.
Wendt, C. 2014. Changing healthcare system types. Social Policy & Administration, 48(7), 864-882. doi.10.1111/spol.12061.
WHO. 2019. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/tobacco.
WHO. 2000. The world health report 2000. Health systems: improving performance. World Health Organization, Geneva.
Woolf, S., Aron, L. (ed.). 2013. US health in international perspective: shorter lives,poorer health. The National Academies/Institute of Medicine, Washington. Retrieved from: https://www.nesri.org/sites/default/files/US_Health_in_International_Perspective.pdf .
Yaya, S., Danhoundo, G. 2015. Introduction: Special issue on innovations in health care system reform in OECD countries. The Public Sector Innovation Journal , 20(1), 1-12.