World Medical & Health Policy

WILEY

RESEARCH ARTICLE

Global Efficiency Assessment in COVID-19 Contagion Control and Treatment

Faruk Yılmaz[1] | İlhan Kerem Şenel[2] | Özgür İnce[2]

[1]Department of Health Management, Faculty of Health Sciences, Muş Alparslan University, Güzeltepe, Muş, Turkey | [2]Department of Health Management, İstanbul University, Cerrahpaşa, İstanbul, Turkey

Correspondence: Faruk Yılmaz (f.yilmaz@alparslan.edu.tr)

Received: 30 January 2025 | **Revised:** 4 May 2025 | **Accepted:** 12 May 2025

Funding: The authors received no specific funding for this work.

Keywords: comparative analysis | COVID-19 | data envelopment analysis | efficiency | machine learning

ABSTRACT

In the global response to the COVID-19 pandemic, countries aimed to reduce morbidity and mortality by enhancing contagion control and treatment efficiency. Efficiency levels varied significantly across countries due to differences in resources, intervention strategies, and contextual factors. This study analyzes 144 WHO member countries to classify them based on preparedness and resources, measure efficiency in contagion control and treatment, and identify determinants of efficiency. A three-stage approach was applied. First, countries were clustered using the k-means algorithm based on development-related indicators, including the Human Development Index, Statistical Performance Index, and Global Health Security Index. Second, contagion control and treatment efficiency were assessed using Data Envelopment Analysis. For contagion control efficiency, input variables included non-pharmaceutical and pharmaceutical interventions, with cases per million as the output. Treatment efficiency was evaluated using health expenditure, hospital beds, and health workforce as inputs, and case fatality rate as the output. Finally, Classification and Regression Trees and Random Forest algorithms were used to determine variables influencing efficiency. The clustering process resulted in two country groups. Findings reveal that health expenditure, urban population percentage, and GDP per capita were significant variables for contagion control efficiency in the first cluster, while testing rates, health expenditure, and GDP per capita were important in the second cluster. For treatment efficiency, land borders, vaccination rates, and GDP per capita emerged as key influencing factors. These findings identify key drivers of efficiency and offer practical guidance for countries to enhance their preparedness and adapt strategies for future global health crises.

1 | Introduction

The COVID-19 pandemic, which emerged in Wuhan, China, on December 31, 2019, rapidly escalated into one of the most significant global health crises of the 21st century. The World Health Organization (WHO) officially declared COVID-19 a pandemic on March 11, 2020, emphasizing its status as an international public health emergency (World Health Organization 2020). By 2023, with over 657 million confirmed cases and 6.67 million deaths worldwide, it had put unprecedented pressure on healthcare systems, emphasizing the necessity for global cooperation. Figures 1 and 2, illustrating COVID-19's spread across WHO regions and World Bank income groups, underscore the varied impacts of the pandemic on different regions and economies (Our World in Data 2023; Yilmaz 2023).

The widespread impact of COVID-19 has underscored the importance of resilient healthcare systems in mitigating the effects of infectious diseases. Throughout the pandemic, healthcare systems faced challenges in implementing public

Summary

- This comprehensive study evaluates COVID-19 efficiency across 144 countries over 2 years.
- A novel three-stage approach assesses COVID-19 contagion control and treatment efficiency.
- Health expenditure and testing frequency are key determinants of contagion control efficiency.
- Border length, vaccination rates, and GDP per capita are key determinants of treatment efficiency.

health measures and managing limited resources (Hale et al. 2021). Both non-pharmaceutical interventions (NPIs), such as quarantine, social distancing, and hygiene protocols, and pharmaceutical interventions, including vaccination and treatment, were employed to varying degrees across countries, affecting the outcomes of pandemic response strategies (Ferguson et al. 2020; Pan et al. 2020). Figure 3 shows the global monthly trends in NPIs and vaccination rates, summarizing these efforts (Our World in Data 2023; Yilmaz 2023).

The immense pressure on healthcare systems by the COVID-19 pandemic has forced many countries to restructure their health services and optimize resource use. The pandemic highlighted the importance of resilient healthcare systems, efficient resource management, and effective public health measures (Ranney et al. 2020). Studies indicate that national responses significantly influence health outcomes and overall healthcare system performance (Anderson et al. 2020; Lotfi et al. 2020). High-income countries generally benefited from greater access to healthcare and quicker responses, while low-income and middle-income countries faced resource constraints that worsened the crisis (Walker et al. 2020).

Preventive and curative health services play a crucial role in achieving one of the fundamental goals of healthcare systems, which is improving overall health outcomes. However, strategies for achieving this goal vary widely, as healthcare systems differ in their structure, financing, and intervention components (Anderson et al. 2020; Legido-Quigley et al. 2020). The success of intervention strategies during health crises is closely tied to healthcare systems' adaptability, preparedness, and resource management capacity. Variations in both NPIs and pharmaceutical interventions directly influenced pandemic management outcomes across countries (Ferguson et al. 2020; Ranney et al. 2020).

In this context, this study proposes a comprehensive framework for assessing the efficiency of countries' COVID-19 management strategies, encompassing both contagion control and treatment performance. Unlike prior studies that focused narrowly on case or death counts or applied single-method approaches, the present research employs an integrated three-stage methodology: clustering, efficiency analysis, and supervised machine learning (ML). This design enables a holistic and context-sensitive examination of national responses.

A key innovation of the study is the preliminary clustering stage, which uses k-means based on the Human Development Index (HDI), Global Health Security (GHS) sub-index, and Statistical Performance Index (SPI) to account for cross-country heterogeneity. This multidimensional and expert-informed grouping enhances the comparability and analytical validity of subsequent efficiency evaluations. Efficiency is then assessed via dual Data Envelopment Analysis (DEA) models—one for contagion control and another for treatment—reflecting the functional distinctions within health systems. Each model incorporates contextually relevant inputs and undesirable outputs, such as total cases and case fatality rate.

Finally, supervised ML algorithms (Classification and Regression Trees [CART] and Random Forest [RF]) are applied to identify the key drivers of efficiency. By integrating DEA with

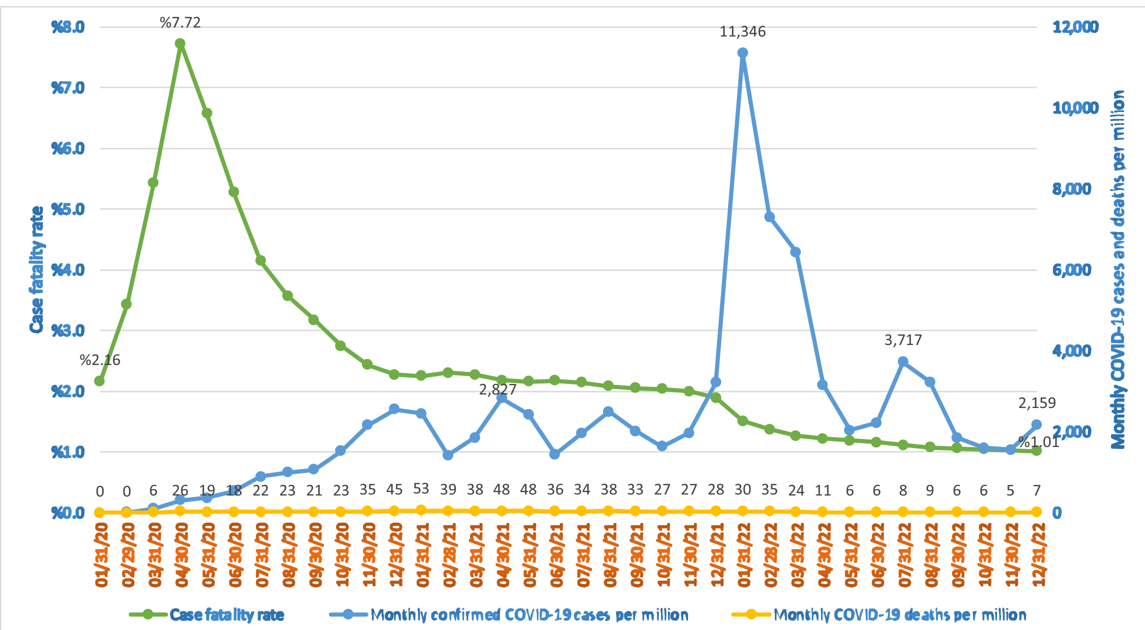

FIGURE 1 | Global impact of COVID-19 (2020–2022).

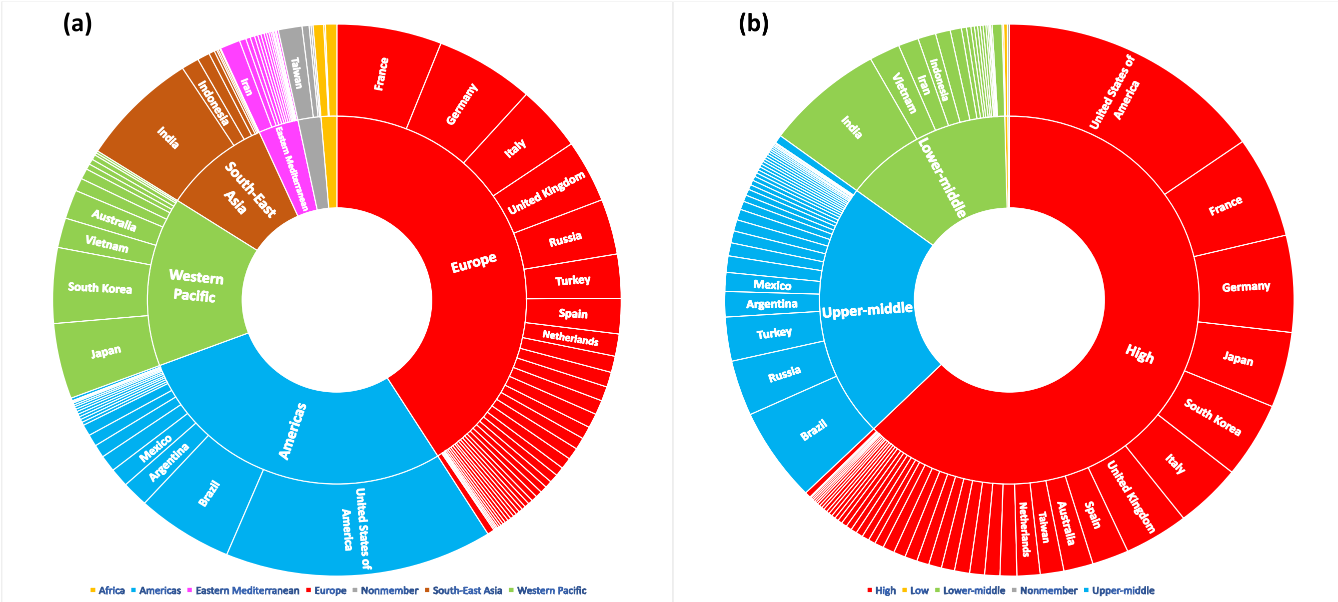

FIGURE 2 | Total COVID-19 cases by WHO region and WB income group.

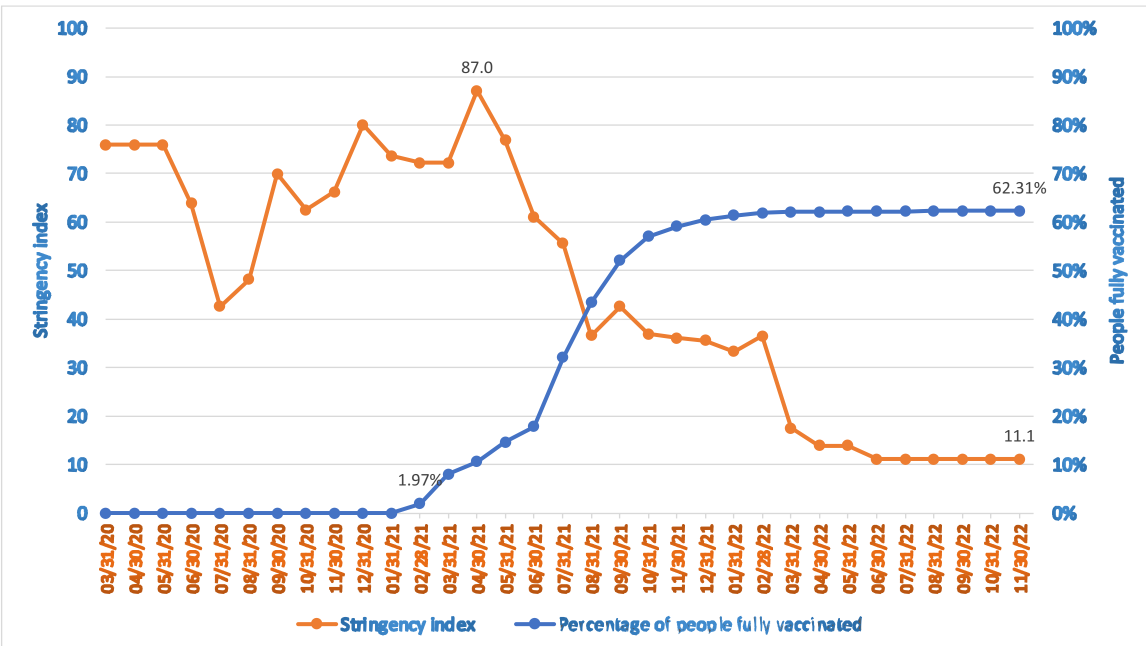

FIGURE 3 | Global monthly trends in stringency index and vaccination rates.

explanatory modeling, this study enhances the interpretability and policy relevance of its findings. In contrast to prior work that used DEA without deeper causal exploration (Breitenbach et al. 2021; Ghasemi et al. 2020; Ordu et al. 2021; Su et al. 2021), this study provides a more nuanced and actionable understanding of efficiency determinants. It contributes critical insights for designing more effective and equitable public health strategies for future global health emergencies.

2 | Literature Review

Assessing healthcare system efficiency remains a high-priority global issue due to its implications for public health outcomes and policy effectiveness. Efficiency encompasses both technical efficiency—achieving greater outputs from given inputs—and allocative efficiency—optimally distributing resources to maximize outcomes at minimal cost (Lupu and Tiganasu 2022; Ozcan 2014). For assessing these dimensions in healthcare systems, particularly in COVID-19 management, sophisticated methodologies that can accommodate multiple input-output relationships are necessary (Breitenbach et al. 2021).

DEA, a non-parametric method to evaluate the relative efficiency of decision-making units (DMUs), is well-suited for such tasks due to its ability to handle multiple inputs and outputs without pre-defined weights. Several studies have used DEA in the context of COVID-19 by analyzing inputs such as healthcare resources (e.g., hospital beds, medical staff), demographic variables (e.g., population density, age), and epidemiological outcomes (e.g., confirmed cases, deaths, recoveries) (Ghasemi et al. 2020; Pereira et al. 2022; Zhu et al. 2022).

More recent research has expanded traditional DEA models to include undesirable outputs and advanced formulations such as super-efficiency or network models (Martínez-Córdoba et al. 2021; Zhu et al. 2022; Pereira et al. 2022; Taherinezhad and Alinezhad 2022). These approaches improve accuracy by reflecting both the successes and failures of pandemic management, especially in capturing undesirable outputs like mortality and infection rates.

ML techniques, such as decision trees, RFs, and ensemble models, have also gained attention for identifying patterns, making predictions, and explaining country-level differences in pandemic performance (Aydin and Yurdakul 2020; Taherinezhad and Alinezhad 2022). ML also facilitates the classification of countries with similar profiles, supporting more tailored and actionable policy guidance. Despite the strengths of DEA and ML, their combined use in a multi-stage framework remains rare. While some studies (Aydin and Yurdakul 2020; Aktas et al. 2022) have applied clustering before DEA, most lack a fully integrated structure combining clustering, DEA, and supervised ML. Additionally, most research has limited temporal scope and single-stage designs, reducing their ability to capture the evolving dynamics of pandemic responses (Ghasemi et al. 2020; Ordu et al. 2021; Su et al. 2021).

To address these gaps, this study proposes a novel three-stage framework that combining k-means clustering, dual DEA models (for contagion control and treatment, both incorporating undesirable outputs), and supervised ML algorithms (decision trees and RF). Unlike previous studies that utilize these methods in isolation or partial combinations, this approach enables a more comprehensive and multidimensional evaluation of country-level pandemic performance. With a data coverage period exceeding 2 years and a structured, multi-stage design, this study represents, to the best of our knowledge, the first to implement such a fully integrated methodology for assessing national COVID-19 responses. Table 1 summarizes previous studies and illustrates the unique methodological contribution of this study.

3 | Methods

3.1 | Research Context

The global management of COVID-19 highlighted the need to assess health systems' preparedness and responses in both contagion control and treatment. Comprehensive evaluations with updated data are crucial for improving pandemic preparedness. This study analyzes the efficiency of 144 WHO member countries regarding COVID-19 morbidity and mortality based on accessible data.

3.2 | Research Design

A three-stage methodology was adopted: (i) forming comparable groups using k-means clustering, (ii) assessing contagion control and treatment efficiency via DEA, and (iii) identifying efficiency determinants through Classification and Regression Tree (CART) and RF algorithms. Figure 4 summarizes the methods, key variables, and steps.

In the clustering stage, three attributes were used to represent development and data quality: the HDI for overall development, the health system sub-score of the GHS Index for pandemic readiness, and the Statistical Performance Indicators (SPI) Index for data reliability. These indicators enabled clustering 144 countries into homogeneous groups. Detailed descriptions are provided in Table 2.

In the second stage, we developed two models to separately evaluate health systems' contagion control and treatment efficiency. These models provide a comprehensive assessment of health system performance. Efficiency was measured using DEA, with adjustments made for undesirable outputs. Inputs and outputs used in each model are presented in Table 3.

In the final stage, countries were classified as "efficient" or "not" based on cutoff points derived from the histograms of contagion control and treatment efficiency scores obtained through DEA. CART and RF algorithms were then used to identify factors contributing to these classifications. Variables selected for analysis were based on prior literature and expert opinions, capturing relevant healthcare, economic, and social dimensions.

The stringency index for January 30, 2020–December 31, 2020 (SI20) and the number of fully vaccinated individuals as of December 31, 2021 (FV21) were included to assess the impact of past interventions during peak infection periods, considering the cross-sectional design of the study. Additionally, the following indicators were analyzed (Our World in Data 2023; World Bank 2023; World Health Organization 2023):

- *Health resource indicators:* Physicians per 1000 people (P), Nurses and midwives per 1000 people (NM), Hospital beds per 1000 people (HB), and Total tests per 1000 people as of May 31, 2022 (T).
- *Health expenditure indicators:* Current health expenditure per capita current US$ (CHEPC), CHE% of GDP (CHEP), domestic general government health expenditure as % of general government expenditure (DGGE), domestic general government health expenditure as % of CHE (DCHE), out-of-pocket expenditure as % of CHE (OOPS).
- *Economic development and public share indicators:* Gross domestic product per capita PPP, $ (GDP), Gini index (GI), unemployment % of total labor (U), and general government expenditure as % GDP (GGE).
- *Education level and public role indicators:* Children out of school % of primary school age (COS), Literacy rate (LR) and Government expenditure on education % of GDP (GEE).
- *Socio-demographic structure indicators:* Population ages 65 and above (% of total population [P65], population density-people per sq. km [PD], and urban population [UP]).
- *Geographic characteristics indicators:* No. of distinct land neighbors (NDLN) and total land borders (LB).

TABLE 1 | A comprehensive literature review of DEA and analytical techniques in global COVID-19 efficiency.

Study	Countries	Data period	Stage 1: Clustering	Stage 2: DEA model	Stage 3: ML and stats tools
Su et al. (2021)	23 countries (G20 + other countries)	105 days post-first confirmed case	Not applied	CCR (CRS) model	Not applied
Ghasemi et al. (2020)	19 countries	February 2–April 12, 2020 (weekly)	Not applied	Dynamic DEA; two separate models (preventing spread and deaths)	Not applied
Ordu et al. (2021)	16 countries (5 weeks post-100th case)	December 31, 2019–April 11, 2020	Not applied	Super-efficiency DEA (80 models)	Not applied
Ibrahim et al. (2020)	58 countries	Up to July 15, 2020	Not applied	CRS model with undesirable outputs; two separate models (contagion control and treatment)	Not applied
Aydin and Yurdakul (2020)	142 countries	January 21–July 28, 2020	k-means and hierarchical clustering	WSIDEA model	Decision tree and random forest
Breitenbach et al. (2021)	36 countries (90% of global cases and deaths)	Up to November 11, 2020	Not applied	BCC (VRS) with You and Yan (2011) ratio model	Not applied
Pereira et al. (2022)	55 countries (OECD + key partners)	Up to 31 December, 2020	Not applied	Network DEA; integrated stages (population, contagion, triage, hospitalization, and ICU admission)	Not applied
Martínez-Córdoba et al. (2021)	155 countries	Up to February 1, 2021	Not applied	BCC (VRS) with Seiford and Zhu (2002) model for undesirable outputs	Truncated regression
Delis et al. (2021)	81 countries	April 1, 2020–March 18, 2021 (monthly)	Not applied	CCR (CRS) model	Bivariate regressions
Taherinezhad and Alinezhad (2022)	100 countries with the highest cases	Up to June 21, 2021	Not applied	Network VRS DEA; integrated systems (prevention, infection detection, and medical)	Bagging, AdaBoost, and RUSBoost
Lupu and Tiganasu (2022)	31 European countries	January 1–December 31, 2020	Not applied	CCR (CRS) model	Tobit regression
Zhu et al. (2022)	117 UN countries	January 20, 2020–August 1, 2021	Not applied	Multi-stage super-efficiency DEA	Not applied
Aktas et al. (2022)	107 countries	June–December 2020	Model-based clustering (3 groups)	DEA with assurance region (AR) constraints	Tobit regression
This study	144 WHO member countries	Up to June 30, 2022	k-means clustering	BCC (VRS) with Seiford and Zhu (2002) model for undesirable outputs; two separate models (contagion control and treatment)	Decision tree and random forest

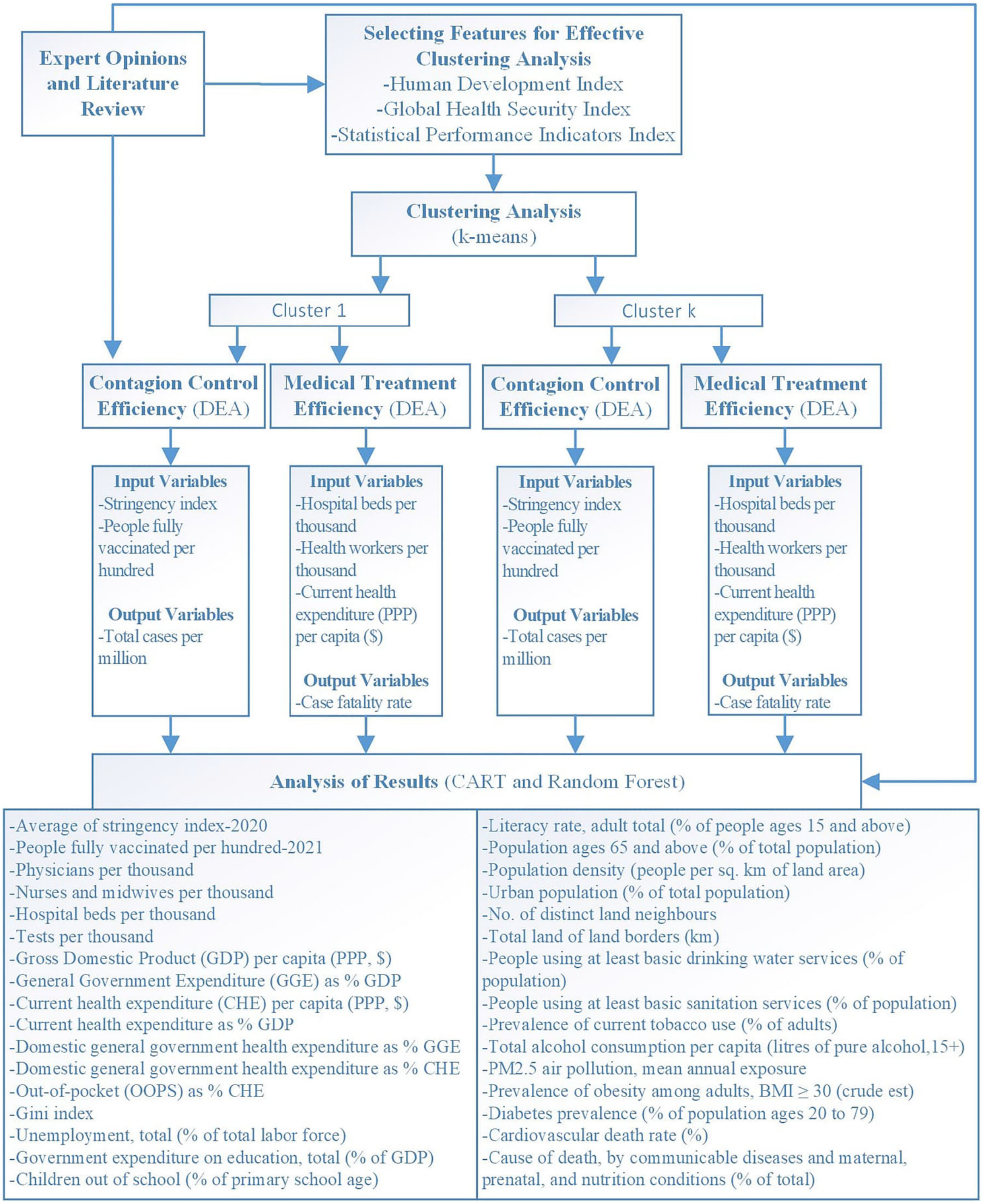

FIGURE 4 | Research design.

- *Health risk and behavioral factor indicators:* People using at least basic drinking water services (DWS), people using at least basic sanitation services (BSS), prevalence of tobacco use (TU), total alcohol consumption per capita (TAC), PM2.5 air pollution (AP), prevalence of obesity among adults, BMI ≥ 30 (OP), and diabetes prevalence (DP).
- *Health status indicators:* Cardiovascular death rate (CDR) and cause of death by communicable diseases and maternal, prenatal, and nutritional conditions (CDD).

Data were sourced from the World Health Organization, World Bank, and Our World in Data. Analyses and visualizations were conducted using R Studio (with "factoextra," "ggplot2," "NbClust," "cluster," and "deaR" packages) and KNIME, following the study's multi-stage analytical framework. This section outlines the analytical models and methods applied in each stage.

3.3 | Stage 1: K-Means Clustering Algorithm

As an unsupervised learning method, k-means partitions the data set into k clusters based on similarities, measured by distance metrics (Macqueen 1967; Kohad 2021). The simplest

TABLE 2 | Indicators used in the clustering analysis.

Indicator name	Calculation methodology	Source and data year
Global Health Security Index—Health System Subscore	Calculated by jointly evaluating the variables across seven subcategories under the health system sub-dimension of the Global Health Security (GHS) Index.	Cameron et al. (2019)
Human Development Index Score	Calculated by taking the geometric mean of the scores for three indices: Life Expectancy Index, Education Index, and Income Index.	United Nations: Human Development Index (HDI) (2020)
Statistical Performance Indicators Index Score	Calculated by taking the weighted average of 22 dimensions under five categories: Data Usage, Data Services Trusted by Users, Data Products, Data Sources, and Data Infrastructure.	World Bank (2023)

TABLE 3 | Input and output variables in the contagion control and treatment efficiency models.

Efficiency model	Variable types	Variable name	Description	Data source
Contagion control efficiency	Input variable	Stringency index (2022)	Calculated by the authors by taking the arithmetic mean of the daily values (30 January, 2020–30 June, 2022) of the stringency index	Ourworld (2022)
Contagion control efficiency	Input variable	People fully vaccinated per hundred	Calculated as the number of people who received at least two doses of vaccine as of June 30, 2022 divided by the total population and multiplied by 100	Ourworld (2022)
Contagion control efficiency	Output variable	Total cases per million	The number of confirmed COVID-19 cases as of June 30, 2022 divided by the population and multiplied by one million	Ourworld (2022)
Medical treatment efficiency	Input variable	Health workers per thousand	Calculated as the total number of physicians, nurses, and midwives divided by the population and multiplied by 1000	World Bank (2013)
Medical treatment efficiency	Input variable	Hospital beds per thousand	The number of patient beds in public, private, general, and specialized hospitals and rehabilitation centers is divided by population and multiplied by 1000	World Bank (2013)
Medical treatment efficiency	Input variable	Current health expenditure (PPP) per capita ($)	Current health expenditures per capita expressed in international dollars at purchasing power parity	World Health Organization (2020)
Medical treatment efficiency	Output variable	Case fatality rate	Calculated by the authors as deaths due to COVID-19 divided by the number of confirmed COVID-19 cases as of June 30, 2022	Ourworld (2022)

method to determine the number of clusters is to use the formula "$k = \sqrt{n/2}$", where n is the number of observations. However, for larger datasets, various indices recommended in the literature are used, including the Elbow Method, Silhouette Index, Dunn Index, and Calinski-Harabasz Index (Rodriguez et al. 2019; Zhou 2021). Since no one index consistently yields the best number of clusters, all were considered alongside expert knowledge and the specific aims of this study.

3.4 | Stage 2: DEA Model With Linear Monotonic Transformation for Undesirable Outputs

DEA is a widely used non-parametric method for evaluating health system efficiency. The original DEA model by Charnes et al. (1978) assumed constant returns to scale (CRS), while the BCC model by Banker et al. (1984) introduced variable returns to scale (VRS), allowing assessment of technical, pure technical, and scale efficiency. Over time, advanced DEA variants have addressed super-efficiency, cost efficiency, the Malmquist Index, and models including undesirable outputs (Ozcan 2014).

In health applications, DEA often encounters undesirable outputs like mortality rates. Conventional models inadequately capture these, leading to the development of methods such as treating them as inputs or excluding them—both of which distort the production process (Jahanshahloo et al. 2005). To address this, Seiford and Zhu (2002) proposed a linear monotonic transformation that retains convexity by converting undesirable outputs into positive values through additive inverse transformation, enabling accurate efficiency measurement.

This study employs an output-oriented VRS model with such a transformation to minimize undesirable outputs while maximizing desirable ones. A score of "1" indicates full efficiency, while scores above "1" represent deviations from efficiency. A key limitation of this model is that it can only be applied under the assumption of variable returns to scale (Hua and Bian 2007).

3.5 | Stage 3: Identification of Efficiency Determinants

In the final stage, factors influencing countries' efficiency classifications were analyzed using ML algorithms. Two supervised ML algorithms were applied.

3.5.1 | Decision Trees–CART

The CART algorithm, developed by Breiman et al. (1984), is a non-parametric decision tree technique used for both classification and regression. Unlike models assuming data distribution (e.g., Naïve Bayes), CART splits data based on feature values using the Gini index to optimize purity at each node (Kelleher et al. 2020; Sullivan 2017). It builds binary trees through recursive partitioning and uses cost-complexity pruning to prevent overfitting (Maimon and Rokach 2014). CART's strengths include its ability to handle missing/continuous data and produce interpretable models (Singh and Giri 2014; Sullivan 2017).

3.5.2 | RF Algorithm

The RF algorithm, developed by Breiman (2001), is an ensemble learning method based on decision trees that achieves high predictive performance by combining bootstrap sampling with a random subspace approach (Rokach 2016). During training, approximately 63.2% of the data is used for building each tree, while the remaining 36.8% serves as a validation set, allowing for estimation of generalization error through in-bag and out-of-bag samples (Zhou 2021). Unlike traditional decision trees, RF selects splitting attributes from randomly chosen subsets of features, with the subset size typically determined as $n = \log_2 N + 1$, based on depending on the number of total features (Breiman 2001; Zhou 2021). The algorithm's main parameters are the number of features considered at each split and the number of trees, with 100 trees generally providing robust performance (Hartshorn 2016). Model performance is evaluated using metrics such as accuracy, sensitivity, specificity, *F*-score, and AUC, and feature importance is assessed via permutation and Gini importance measures (Hjerpe 2016; Zhou 2021). Due to its ability to deliver high accuracy and fast learning across diverse datasets, RF remains a highly effective classification method (Fernandez-Delgado et al. 2014).

4 | Results

4.1 | Clustering Analysis Findings

In the initial stage, 144 countries were classified into homogeneous subgroups using the k-means algorithm, a common unsupervised ML method. The optimal number of clusters was determined as two, based on the majority rule from 11 indices in the "NbClust" package in R Studio. Based on this assignment, 73 countries were grouped into Cluster 1, while 71 countries formed Cluster 2. A two-dimensional representation of the cluster allocation, generated through the "factoextra" package in R Studio, is presented in Figure 5.

Figure 5 illustrates that Cluster 1 comprises 73 countries, including the United States, Germany, Australia, China, France, India, the United Kingdom, Spain, and Turkey. Conversely, Cluster 2 encompasses 71 countries, such as Afghanistan, Bangladesh, Morocco, South Africa, Egypt, Uzbekistan, Pakistan, Vietnam, and Zimbabwe. The total within-cluster sum of squares was computed as 74.545 for Cluster 1 and 86.666 for Cluster 2, with an intercluster variance ratio of 62.4% relative to the total data set variance. This clustering outcome demonstrates an effective grouping of countries with similar characteristics.

4.2 | DEA Findings

Following clustering, the DEA specifically an output-oriented VRS model, for undesirable outputs as developed by Seiford and Zhu (2002), was applied to assess each country's efficiency in COVID-19 contagion control and treatment. This analysis was conducted using an output-oriented VRS model, with input and output variables normalized. Variable normalization was performed, dividing each variable's value by its group arithmetic mean. The findings from the contagion control and treatment efficiency models applied in R Studio are presented in two main sections.

4.2.1 | DEA Findings for Contagion Control Efficiency

For contagion control efficiency, input variables were defined as the mean value of the stringency index and the percentage of fully vaccinated individuals, while the output variable was the total number of cases per million. After normalizing these variables, necessary definitions were made using the "make_deadata" function from the "deaR" package, and analyses were conducted for both clusters. A histogram distribution of the contagion control efficiency scores for countries in Clusters 1 and 2, analyzed in R, is presented in Figure 6.

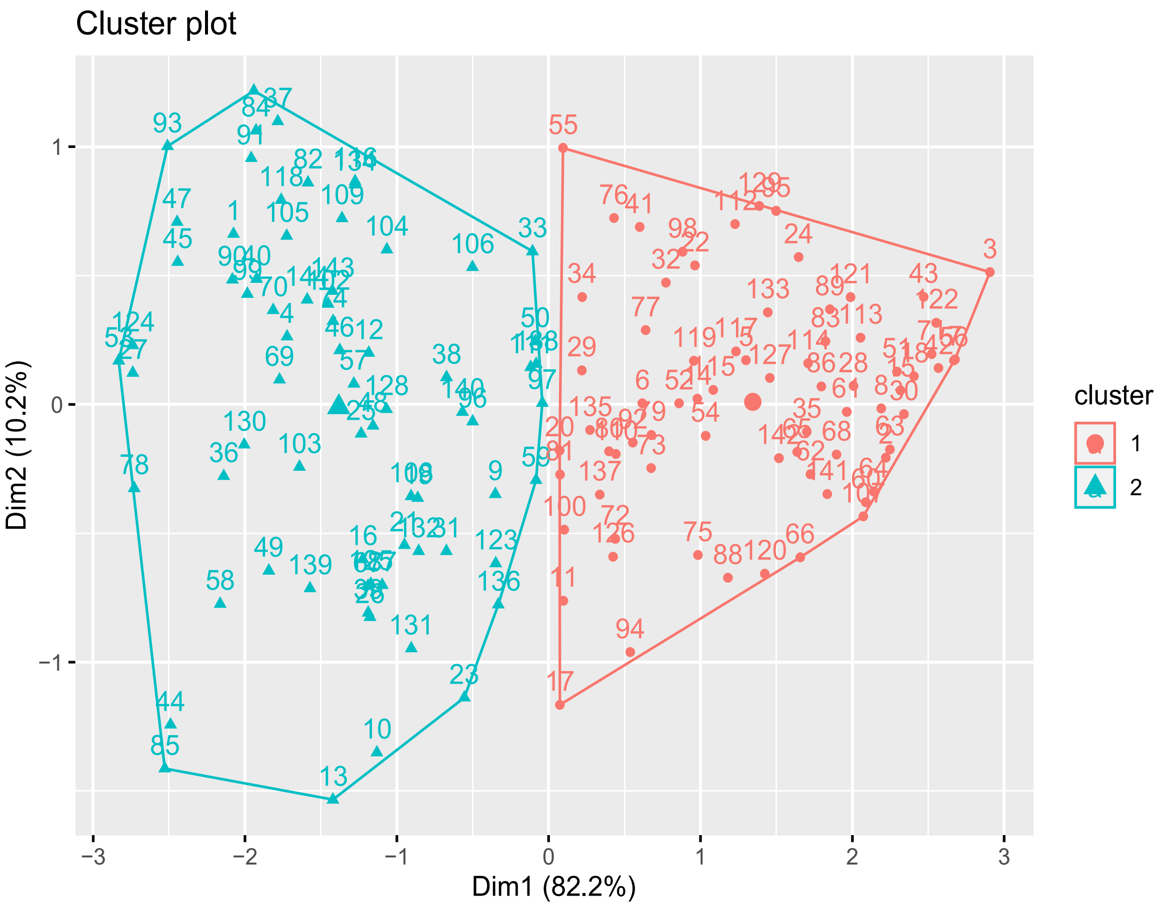

FIGURE 5 | Clustering results using the k-means algorithm.

Identifying appropriate class labels for efficiency determinants involved applying a histogram distribution analysis, a method frequently referenced in the literature. Based on this distribution, the potential cutoff point for efficiency scores was set at 1.0002. Countries with scores between "$1 \leq ES < 1.0002$" were classified as "efficient," while those with "$ES \geq 1.0002$" were categorized as "not". In Cluster 1, 15 countries met the criteria for contagion control efficiency, while 58 were identified as "not". In Cluster 2, 32 countries were classified as "efficient" and 39 as "not". Contagion control and treatment efficiency scores and classes for Cluster 1 are presented in Table 4.

4.2.2 | DEA Findings for Treatment Efficiency

For treatment efficiency, input variables were health workers per thousand, hospital beds per thousand, and healthcare expenditure per capita, with the case fatality rate as the output. Similar to previous analyses, normalization was performed, and the DEA process was carried out using the "deaR" package. Cluster 1 included 27 countries classified as "efficient" and 46 as "not". Contagion control efficiency scores and classes for Cluster 2 are provided in Table 5.

Conversely, the treatment efficiency findings for countries in Cluster 2 did not yield meaningful results, primarily due to insufficient health resources to fulfill basic healthcare functions. As indicated by the World Health Organization (2016), a minimum density of 2.3 health workers (including physicians, nurses, and midwives) per 1000 individuals is essential to achieve adequate primary healthcare coverage. Nevertheless, in approximately 50% of the countries within Cluster 2, health worker density was below this threshold. This insufficiency reflects a constraint in the capacity to deliver primary healthcare services, which, in turn, limits the meaningful evaluation of treatment efficiency for these countries. Cluster 1 countries, however, were assessed on both contagion control and treatment efficiency. The classification of these countries by efficiency type is detailed in Table 6.

As shown in Table 6, out of the 73 countries in Cluster 1, five countries—Belarus, China, India, Japan, and Thailand (6.85%)—were fully efficient. Ten countries (13.70%) were efficient only in contagion control, 22 countries (30.14%) were efficient only in treatment, and 36 countries (49.31%) were inefficient in both dimensions. These results will be further analyzed to identify factors explaining countries' efficiency in COVID-19 contagion control and treatment.

4.3 | CART and RF Algorithm Findings

In the final stage, CART and RF algorithms identified factors underlying contagion control and treatment efficiency. Conducted in KNIME, the analysis included four sub-analyses for contagion control efficiency in both clusters and two

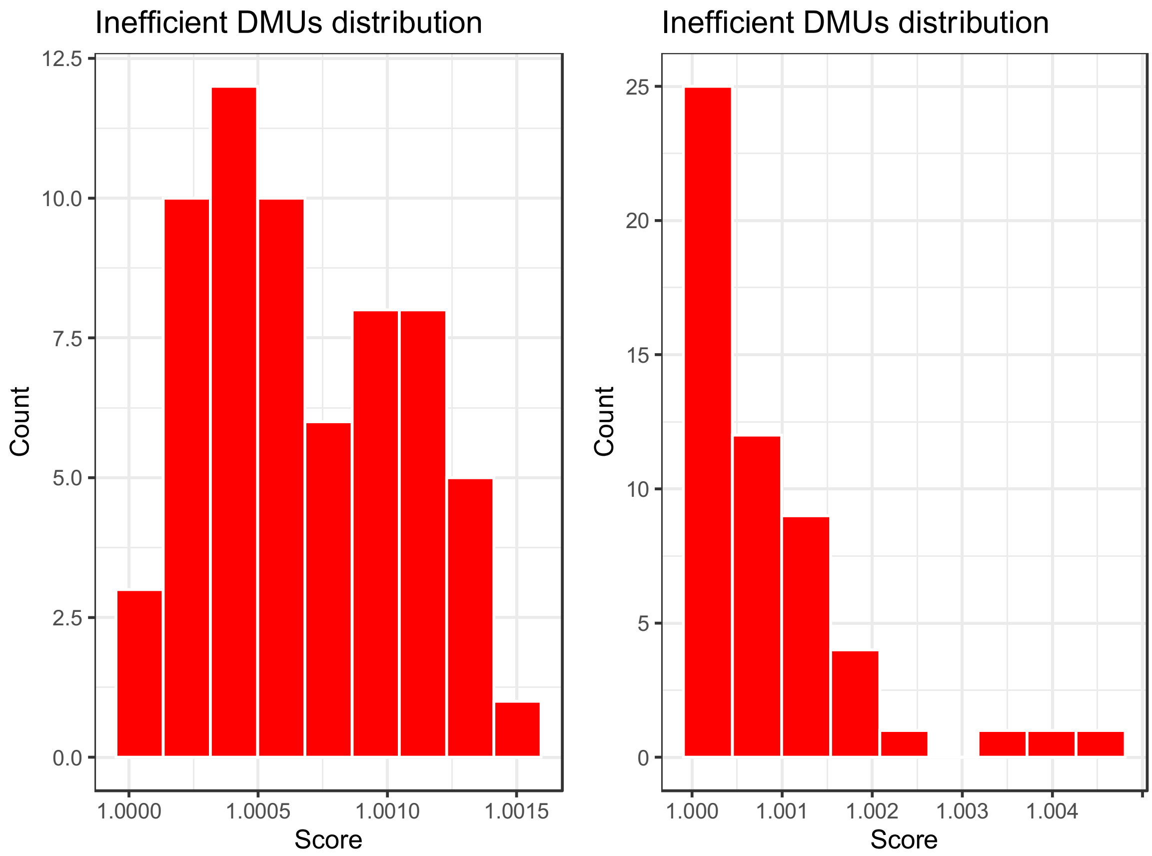

FIGURE 6 | Histogram of contagion control efficiency scores for Clusters 1 and 2, respectively.

sub-analyses for treatment efficiency. Additionally, RF results obtained in R Studio were presented to determine the importance of explanatory variables. The data set was divided into training (70%) and test (30%) sets using stratified sampling for supervised learning problems, with the exception of contagion control analysis in Cluster 1, where these proportions were adjusted to 66.7% and 33.3%, respectively, to address the class imbalance. Decision trees derived from the Gini index-based CART model are displayed in Figure 7.

In examining contagion control efficiency within Cluster 1, healthcare expenditure per capita (CHEPC) emerged as the primary explanatory factor. Among 37 countries with CHEPC exceeding USD 1523.22, only Japan was classified as efficient, while the remaining 36 countries were not. This finding suggests that higher healthcare expenditures alone do not ensure effective containment of COVID-19 and may instead reflect increased spending on case detection. Additionally, the presence of urban population as a sub-node underscores the challenges in densely populated areas, where high contact rates can complicate effective control measures.

In Cluster 2, the number of tests per thousand emerged as the root node, with only Suriname classified as inefficient among the 21 countries with testing rates below 187.5 per thousand. This result implies that low testing rates may lead to underreporting of cases. Furthermore, countries with a GDP per capita below USD 5190 seem to encounter constraints in healthcare accessibility, potentially distorting efficiency assessments in environments with insufficient testing capabilities.

For treatment efficiency in Cluster 1, the land border emerged as the root explanatory variable. This might reflect the relative ease of healthcare management in countries with shorter land borders, such as Estonia, Bahrain, Costa Rica, Singapore, the United Arab Emirates, South Korea, Iceland, the Netherlands, and Portugal, many of which also have compact geographical areas. Shorter borders may enhance patient access, while larger borders demand broader resource allocation for healthcare services, including cross-border demands. In the initial sub-node, countries with high percentages of out-of-pocket healthcare expenditures were inefficient, possibly due to limited financial accessibility to healthcare. Subsequent sub-nodes included additional explanatory variables, such as obesity prevalence and nurses and midwives per thousand.

The study evaluated model performance using receiver operating characteristic (ROC) curves, with a higher area under the curve (AUC) indicating greater predictive accuracy. Figure 8 displays the ROC curves for the CART models used in this study: C1CECART, representing the Contagion Control Efficiency CART Model for Cluster 1; C2CECART, the Contagion Control Efficiency CART Model for Cluster 2; and C1TECART, the Treatment Efficiency CART Model for Cluster 1.

AUC values for the contagion control efficiency prediction model in Cluster 1, Cluster 2, and the treatment efficiency prediction model in Cluster 1 were calculated as 0.816, 0.95, and 0.875, respectively. These AUC values demonstrate the CART models' strong predictive performance.

TABLE 4 | Contagion control and treatment efficiency scores and classes for Cluster 1.

Countries	CES	TES	CEC	TEC	Countries	CES	TES	CEC	TEC
Albania	1.0002	1.0004	Efficient	Not	Kuwait	1.0004	1.0001	Not	Efficient
Argentina	1.0005	1.0007	Not	Not	Kyrgyzstan	1	1.0002	Efficient	Not
Australia	1.0008	1	Not	Efficient	Latvia	1.0012	1.0003	Not	Not
Austria	1.0013	1.0002	Not	Not	Lithuania	1.0011	1.0003	Not	Not
Bahrain	1.0011	1	Not	Efficient	Luxembourg	1.001	1.0002	Not	Efficient
Belarus	1	1.0002	Efficient	Efficient	Malaysia	1.0003	1.0001	Not	Efficient
Belgium	1.0009	1.0004	Not	Not	Malta	1.0005	1.0003	Not	Not
Bosnia	1	1.0022	Efficient	Not	Mauritius	1	1.0012	Efficient	Not
Brazil	1.0004	1.001	Not	Not	Mexico	1.0001	1.0026	Efficient	Not
Bulgaria	1.0002	1.0017	Not	Not	Moldova	1.0003	1.001	Not	Not
Canada	1.0002	1.0006	Not	Not	Netherlands	1.0012	1.0001	Not	Efficient
Chile	1.0005	1.0007	Not	Not	New Zealand	1.0006	1	Not	Efficient
China	1	1	Efficient	Efficient	Norway	1.0006	1.0001	Not	Efficient
Colombia	1.0003	1.001	Not	Not	Panama	1.0005	1.0004	Not	Not
Costa Rica	1.0005	1.0001	Not	Efficient	Peru	1.0003	1.0029	Not	Not
Croatia	1.0006	1.0007	Not	Not	Philippines	1	1.0003	Efficient	Not
Cuba	1.0002	1.0003	Not	Not	Poland	1.0003	1.001	Not	Not
Cyprus	1.0015	1	Not	Efficient	Portugal	1.0013	1.0002	Not	Efficient
Czech Republic	1.001	1.0005	Not	Not	Qatar	1.0003	1	Not	Efficient
Denmark	1.0014	1.0001	Not	Efficient	Romania	1.0003	1.0011	Not	Not
Ecuador	1.0001	1.0018	Efficient	Not	Russia	1.0003	1.0011	Not	Not
Estonia	1.001	1.0001	Not	Efficient	Saudi Arabia	1	1.0005	Efficient	Not
Finland	1.0005	1.0002	Not	Efficient	Serbia	1.0007	1.0003	Not	Not
France	1.0012	1.0002	Not	Not	Singapore	1.0006	1	Not	Efficient
Georgia	1.0011	1.0003	Not	Not	Slovakia	1.0008	1.0005	Not	Not
Germany	1.0009	1.0002	Not	Not	Slovenia	1.0013	1.0003	Not	Not
Greece	1.001	1.0003	Not	Not	Spain	1.0007	1.0004	Not	Not
Hungary	1.0004	1.0013	Not	Not	Sweden	1.0006	1.0004	Not	Not
Iceland	1.0013	1	Not	Efficient	Switzerland	1.0011	1.0002	Not	Efficient
India	1	1	Efficient	Efficient	Thailand	1.0001	1	Efficient	Efficient
Indonesia	1	1.0009	Efficient	Not	Türkiye	1.0004	1.0001	Not	Efficient
Ireland	1.0008	1.0002	Not	Not	Ukraine	1.0003	1.001	Not	Not
Israel	1.0012	1	Not	Efficient	UAE	1.0002	1	Not	Efficient
Italy	1.0008	1.0004	Not	Not	UK	1.0009	1.0004	Not	Not
Japan	1.0001	1.0001	Efficient	Efficient	USA	1.0006	1.0006	Not	Not
Kazakhstan	1.0001	1.0005	Efficient	Not	Uruguay	1.0007	1.0003	Not	Not
Korea Republic	1.0009	1	Not	Efficient					

Abbreviations: CEC, contagion efficiency class; CES, contagion efficiency score; TEC; treatment efficiency class; TES, treatment efficiency score.

This study also used the RF algorithm as an alternative to the CART algorithm. To predict contagion control efficiency, the Contagion Control Efficiency RF Models for Cluster 1 (C1CERF) and Cluster 2 (C2CERF) were developed, where efficiency class was defined as the target variable, the Gini index served as the splitting criterion, and 100 trees were specified for model stability. Additionally, the Treatment Efficiency RF Model for Cluster 1 (C1TERF) was applied to evaluate treatment efficiency, thereby providing a comprehensive framework for assessing both contagion control and treatment outcomes. In classification tasks, KNIME calculates the number of randomly selected input features in each subset based on the square root of the total number of explanatory variables. After splitting the data set into 70% training and 30% test subsets,

TABLE 5 | Contagion control efficiency scores and classes for Cluster 2.

Countries	CES	CEC	Countries	CES	CEC	Countries	CES	CEC
Afghanistan	1	Efficient	Gambia	1	Efficient	Namibia	1.0008	Not
Algeria	1.0001	Efficient	Ghana	1	Efficient	Nepal	1.0004	Not
Angola	1	Efficient	Guatemala	1.0006	Not	Nicaragua	1	Efficient
Azerbaijan	1.001	Not	Guinea	1	Efficient	Nigeria	1	Efficient
Bahamas	1.0011	Not	Guyana	1.0011	Not	Oman	1.0011	Not
Bangladesh	1.0001	Efficient	Haiti	1	Efficient	Pakistan	1.0001	Efficient
Barbados	1.0038	Not	Honduras	1.0005	Not	Paraguay	1.0012	Not
Belize	1.002	Not	Iran	1.001	Not	Rwanda	1.0001	Efficient
Bolivia	1.001	Not	Iraq	1.0007	Not	Senegal	1	Efficient
Botswana	1.0016	Not	Jamaica	1.0006	Not	South Africa	1.0008	Not
Brunei	1.0047	Not	Jordan	1.0019	Not	Sri Lanka	1.0004	Not
Cabo Verde	1.0013	Not	Kenya	1.0001	Efficient	Sudan	1	Efficient
Cambodia	1.0001	Efficient	Lebanon	1.0026	Not	Suriname	1.0017	Not
Cameroon	1	Efficient	Lesotho	1.0002	Efficient	Tajikistan	1	Efficient
Congo	1	Efficient	Liberia	1	Efficient	Timor-Leste	1.0002	Not
Cote d'Ivoire	1	Efficient	Libya	1.0009	Not	Trinidad & Tobago	1.0014	Not
Djibouti	1.0002	Efficient	Madagascar	1	Efficient	Tunisia	1.0011	Not
Dominican Republic	1.0007	Not	Malawi	1	Efficient	Uganda	1	Efficient
Egypt	1	Efficient	Mali	1	Efficient	Uzbekistan	1.0001	Efficient
El Salvador	1.0003	Not	Mauritania	1.0001	Efficient	Venezuela	1.0002	Not
Eswatini	1.0008	Not	Mongolia	1.0035	Not	Vietnam	1.0014	Not
Ethiopia	1	Efficient	Morocco	1.0004	Not	Zambia	1.0002	Efficient
Fiji	1.0009	Not	Mozambique	1.0001	Efficient	Zimbabwe	1.0002	Efficient
Gabon	1.0002	Not	Myanmar	1.0001	Efficient			

Abbreviations: CEC, contagion efficiency class; CES, contagion efficiency score.

TABLE 6 | Efficiency class distribution of countries in Cluster 1.

CCI-TI			CCI-TE		CCE-TI	CCE-TE
USA	Ireland	Panama	Australia	Qatar	Albania	Belarus
Germany	Spain	Peru	Bahrain	Cyprus	Bosnia	China
Argentina	Sweden	Poland	UAE	Costa Rica	Ecuador	India
Austria	Italy	Romania	Denmark	Kuwait	Indonesia	Japan
Belgium	Canada	Russia	Estonia	Luxembourg	Philippines	Thailand
United Kingdom	Colombia	Serbia	Finland	Malaysia	Kazakhstan	
Brazil	Cuba	Slovakia	South Korea	Norway	Kyrgyzstan	
Bulgaria	Latvia	Slovenia	Netherlands	Portugal	Mexico	
Czech Republic	Lithuania	Chile	Israel	Singapore	Mauritius	
France	Hungary	Ukraine	Switzerland	Türkiye	Saudi Arabia	
Georgia	Malta	Uruguay	Iceland	New Zealand		
Croatia	Moldova	Greece				

Abbreviations: CCE, contagious control efficient; CCI, contagious control inefficient; TE, treatment efficient; TI, treatment inefficient.

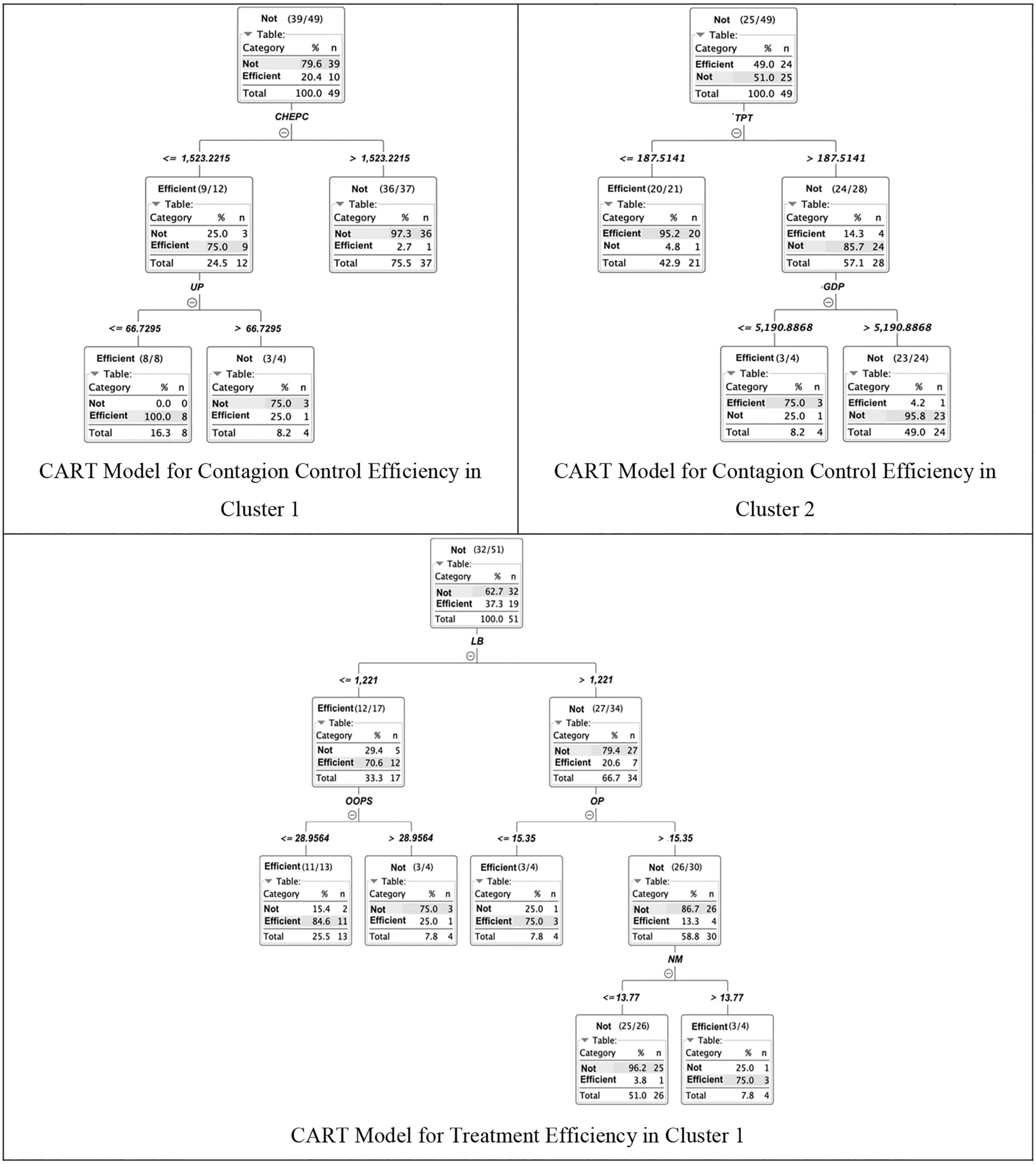

FIGURE 7 | Decision trees of CART models.

out-of-bag (OOB) error rates were used to estimate generalization error and determine variable importance. These findings informed the decision rules used for classifying efficiency in the test set, with majority voting weighted by error rates. Additional model performance metrics, including confusion matrix error rates, ROC curves, and AUC values, were also computed for the test set. In R Studio, similar model optimization procedures identified the 10 most influential variables for each model based on mean decrease in accuracy and Gini purity. Table 7 presents the frequency of top 10 variables in root or primary sub-nodes across training iterations.

Table 7 presents the selection frequency of the top 10 variables for each of the three RF models in the training data. Figure 9 further highlights each variable's importance, measured by its contribution to accuracy and Gini index impurity.

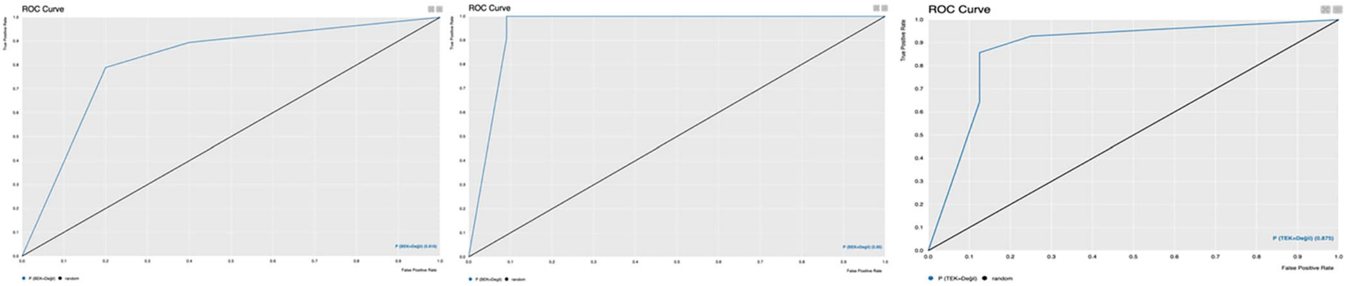

FIGURE 8 | ROC curves for C1CECART, C2CECART, and C1TECART models, respectively.

TABLE 7 | Selection percentage of variables in RF models.

C1CERF			C2CERF			C1TERF		
Variables	**Level 0**	**Level 1**	**Variables**	**Level 0**	**Level 1**	**Variables**	**Level 0**	**Level 1**
CHEPC	83.3%	16.2%	*CHEPC*	90.9%	31.0%	*FV21*	72.2%	25.9%
DCHE	71.4%	13.2%	*TPT*	85.7%	47.4%	*LB*	68.6%	26.5%
UP	66.7%	30.0%	*GDP*	75.0%	28.6%	*GDP*	51.5%	29.4%
GDP	64.7%	19.4%	*DWS*	68.8%	16.0%	*AP*	38.5%	24.2%
TPT	50.0%	33.3%	*FV21*	43.8%	23.3%	*NDLN*	33.3%	19.0%
OOPS	45.5%	21.2%	*LR*	40.0%	33.3%	*PD*	31.4%	22.0%
OP	42.1%	38.2%	*HB*	28.6%	18.2%	*UP*	27.8%	15.6%
DWS	40.0%	21.7%	*NM*	26.3%	26.1%	*NM*	27.3%	26.8%
P	33.3%	2.4%	*OP*	22.2%	3.7%	*OP*	25.0%	31.3%
NM	27.8%	20.0%	*P*	18.8%	14.3%	*COS*	22.9%	14.0%

Both calculation methods identified CHEPC, UP, and TPT as the primary variables for explaining contagion control efficiency in Cluster 1, while TPT, CHEPC, and GDP were key predictors for Cluster 2. For treatment efficiency in Cluster 1, the variables LB, GDP, FV21, PD, and AP emerged as the most important. The ROC curves for the RF models used in this study are displayed in Figure 10.

The AUC values for the contagion control efficiency model were 0.879 in Cluster 1 and 1.0 in Cluster 2, with the treatment efficiency model in Cluster 1 also achieving an AUC of 0.879. In the contagion control efficiency prediction model for Cluster 1, the AUC was calculated as 0.879; for Cluster 2, it reached 1.0; and for the treatment efficiency model in Cluster 1, the AUC was 0.879. These AUC values suggest that RF models offer superior predictive performance in distinguishing efficiency classes compared to CART models. Table 8 presents various performance metrics for the test data set predictions. Overall, both algorithms perform effectively in predicting efficiency classes, though the RF model generally delivers more accurate predictions than the CART model.

5 | Discussion

This study provides a comprehensive efficiency assessment of COVID-19 contagion control and treatment a broad sample of countries, using a novel three-stage methodology that integrates k-means clustering, DEA, and supervised ML. Unlike previous studies that often rely on single-dimensional clustering criteria, such as income, this study integrates multidimensional indicators (HDI, GHS, and SPI) to improve cluster homogeneity and enable more insightful comparisons. To our knowledge, this is the first study to use such a wide-ranging set of attributes for clustering.

Previous studies frequently use single-stage approaches with outputs such as case and death counts (Martínez-Córdoba et al. 2021; Ordu et al. 2021). In contrast, our methodology distinguishes between contagion control and treatment efficiency, utilizing separate models with relevant inputs, such as NPIs and vaccination rates. This dual-dimension analysis is limited in the literature (Ghasemi et al. 2020; Ibrahim et al. 2020). Additionally, the use of supervised ML algorithms for explaining efficiency classes introduces a novel perspective in efficiency evaluation (Aydin and Yurdakul 2020; Xu et al. 2021). This study spans over 2 years of COVID-19 data, including cases, deaths, testing, and vaccinations, enabling a more robust assessment of countries' sustained efficiency in pandemic management.

Through k-means clustering, we grouped 144 countries into two clusters: Cluster 1, which includes advanced economies like the United States, the United Kingdom, Germany, China, Japan, and Türkiye; and Cluster 2, which consists of developing economies like Afghanistan, Bangladesh, and Nigeria. This multidimensional clustering method departs from existing literature. For instance, Carrillo-Larco and Castillo-Cara (2020),

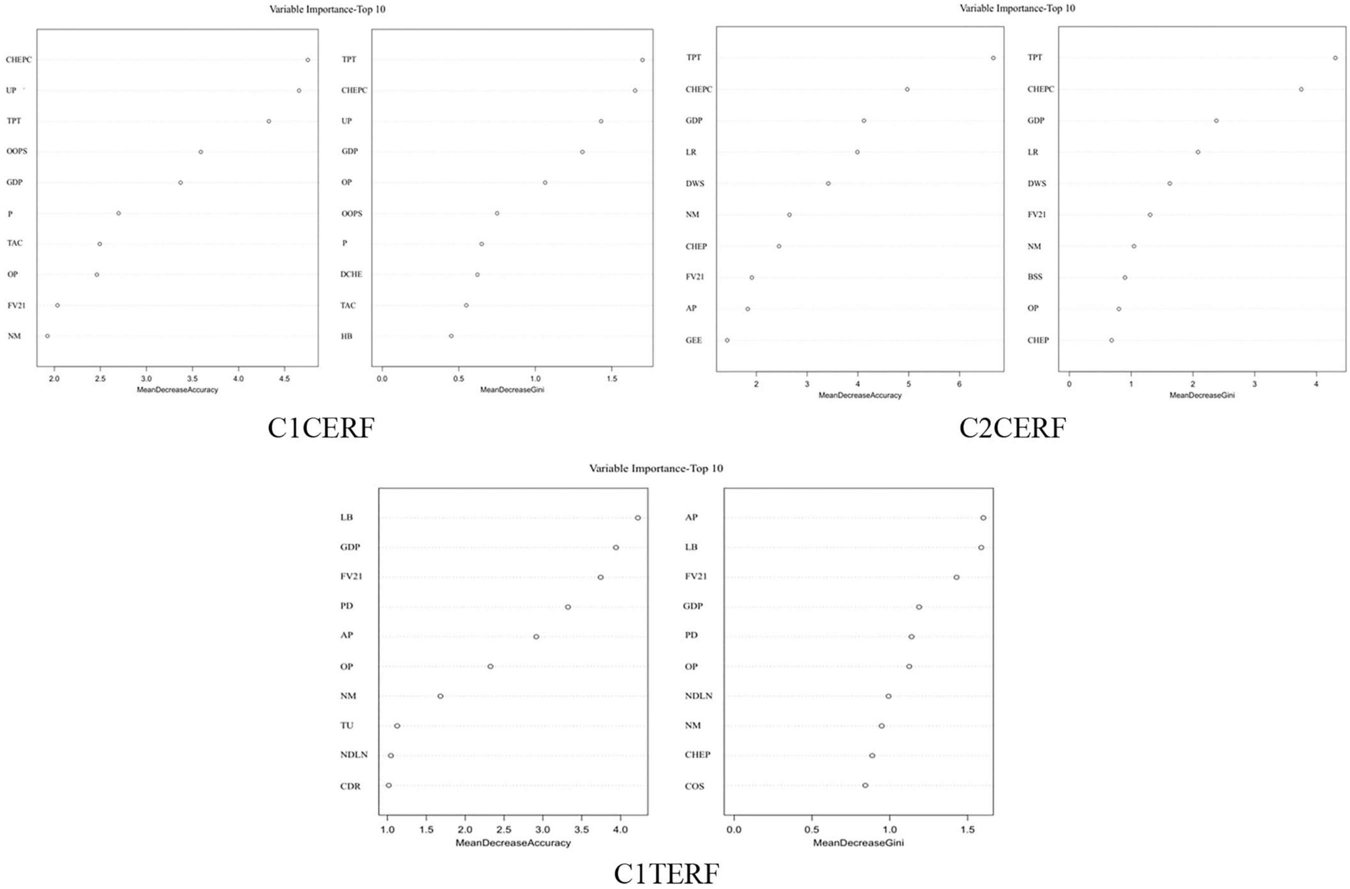

FIGURE 9 | Variable importance levels in RF models based on mean decrease accuracy and Gini.

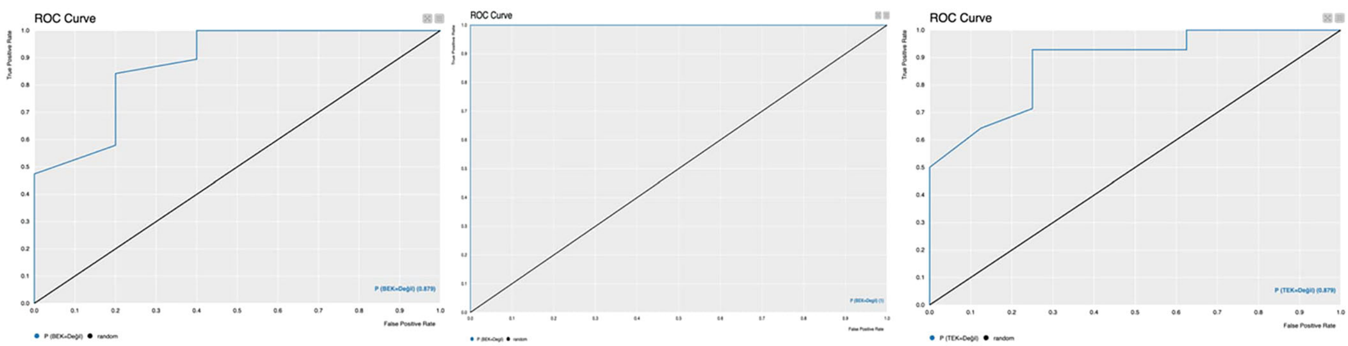

FIGURE 10 | ROC curves for C1CERF, C2CERF, and C1TERF models, respectively.

grouped countries into five clusters based on pre-pandemic health attributes, while other studies (Imtyaz et al. 2020; Rizvi et al. 2021) utilized variables like disease prevalence, socio-economic, and environmental indicators for clustering. Few studies combine clustering and DEA for COVID-19 efficiency evaluation. Aydin and Yurdakul (2020) divided 142 countries into three clusters based on variables such as total cases, total deaths, GDP, and total tests. Aktas et al. (2022) grouped countries by competitive status using the Global Competitiveness Index, categorizing countries as highly competitive, competitive, and noncompetitive. Such comprehensive studies are valuable for a nuanced understanding of COVID-19 efficiency in varied country contexts.

In the second stage, our efficiency assessment of COVID-19 contagion control and treatment aligns with and contrasts findings from prior literature. For instance, Ibrahim et al. (2020) evaluated 58 countries and found that 89.6% were inefficient in contagion control, using population density and International Health Regulations Core Capacity Scores as inputs, with confirmed cases as outputs. Taherinezhad and Alinezhad (2022) identified China, Montenegro, Qatar, Bahrain, and Denmark as fully efficient among 100 countries. Our study's findings align, indicating full efficiency for countries like China, Japan, and Mauritius in Cluster 1, and Nigeria, Sudan, and Mali in Cluster 2, thus providing a broader comparative framework for understanding efficiency variations.

TABLE 8 | Performance metrics for CART and RF models.

	C1CECART	C1CERF	C2CECART	C2CERF	C1TECART	C1TERF
Precision	0.895	0.905	0.909	1	0.923	0.813
Sensitivity	0.895	1	0.909	0.909	0.857	0.929
Specifity	0.6	0.6	0.909	1	0.875	0.625
F-measure	0.895	0.95	0.909	0.952	0.889	0.867
Accuracy	0.833	0.917	0.909	0.955	0.784	0.864
Error	0.167	0.083	0.091	0.045	0.216	0.136
Cohen's kappa	0.495	0.704	0.818	0.909	0.523	0.713
AUC	0.816	0.879	0.95	1	0.875	0.879

In examining treatment efficiency, defined as the success in reducing COVID-19 mortality, output variables such as death rates, recovery rates, and case fatality rates are commonly utilized. For instance, Ibrahim et al. (2020) used inputs such as confirmed cases, physician numbers, and hospital beds, with outputs including COVID-19-related deaths and recoveries, finding that 79% of 58 countries were inefficient in treatment. Afghanistan, Australia, and China demonstrated efficiency, while countries like the United Kingdom and Italy did not. In a separate study analyzing 142 countries in three clusters, Aydin and Yurdakul (2020) found Brunei, Norway, Iceland, Benin, Burundi, Ethiopia, Qatar, Singapore, Uzbekistan, and Venezuela to be efficient in different clusters.

Martínez-Córdoba et al. (2021) analyzed 155 countries with available COVID-19 data, noting that China, Australia, and some African countries were efficient, whereas countries in Europe and the Americas received lower efficiency scores. Lupu and Tiganasu (2022), focusing on the treatment efficiency of European countries in the first wave, found Norway, Finland, and Ireland to be efficient, while Belgium and Germany scored lower. Pereira et al. (2022), using a Network DEA model that included variables such as healthcare spending and PPE usage, showed that Australia, Germany, and New Zealand had high-efficiency scores. Taherinezhad and Alinezhad (2022) also identified China, Qatar, and Denmark as fully efficient.

Our findings similarly indicate full efficiency for countries such as Bahrain, South Korea, Singapore, Iceland, China, New Zealand, Thailand, India, and Qatar. Notably, China, Australia, Estonia, Qatar, and Norway align closely with findings from previous studies (Lupu and Tiganasu 2022; Martínez-Córdoba et al. 2021; Taherinezhad and Alinezhad 2022; Pereira et al. 2022; Ibrahim et al. 2020). Despite variations in model specifications across studies, these findings contribute significant, literature-aligned insights.

In the third stage, we identified key determinants for contagion control efficiency in Cluster 1, such as health expenditure per capita, urban population, and testing rates. For Cluster 2, variables such as GDP per capita and test frequency significantly impacted efficiency classifications. Higher healthcare expenditures and testing frequencies are associated with improved detection, which can potentially lower overall efficiency scores. However, thanks to increased testing, measures such as lockdowns or mask mandates can become more effective, as the identification of cases supports targeted containment efforts. For instance, countries like China, Belarus, and Saudi Arabia have demonstrated efficiency in contagion control, with China and Belarus also achieving high efficiency in treatment outcomes. Additionally, a high case fatality rate is observed in countries where cases may go undiagnosed, indicating gaps in containment efforts.

Studies such as Aydin and Yurdakul (2020) identified number of hospital beds, diabetes prevalence, and stringency index as key variables in explaining efficiency classes. This study focused on treatment efficiency in Cluster 1, where resource constraints were less pronounced. In Cluster 1, the RF algorithm highlighted land border length, GDP per capita, and vaccination rates as significant predictors, aligning with Cheshmehzangi et al. (2021), who found extensive land borders contributed to virus spread. Furthermore, the influence of vaccination rates and GDP per capita is consistent with findings by Xu et al. (2021), who identified these factors as critical for efficiency in US states, and Aktas et al. (2022), who emphasized the role of competitiveness, air pollution, and economic conditions in explaining efficiency. Additional studies (Delis et al. 2023; Lupu and Tiganasu 2022; Martínez-Córdoba et al. 2021) emphasize the role of demographic, environmental, and political factors in managing COVID-19 outcomes.

6 | Limitations

This study is limited to countries with accessible data, excluding those with incomplete information, which narrows the scope of the findings. The comparative analysis was conducted using clusters generated by the k-means algorithm; however, alternative clustering methods may yield different results. The diversity of input-output variables in the second stage may lead to varying findings across studies, underlining the importance of contextual interpretation. The use of the SPI for clustering excluded countries with limited data collection and sharing capacities. However, factors such as the accuracy of COVID-19 data reported by countries, undetected asymptomatic cases, reluctance to seek healthcare, inadequate screening programs, and healthcare resource constraints may affect the reliability of the reported data. As such, the study's findings are based exclusively on data provided by national health authorities.

7 | Conclusions

This study provides a comprehensive assessment of the global response to COVID-19, using ML algorithms to identify key factors that influenced pandemic outcomes. The findings underscore the essential role of preparedness processes and intervention strategies in achieving contagion control and treatment efficiency. Notably, countries with similar resources but more successful outcomes highlighted specific strategies that could serve as effective guides for policymakers.

Our research indicates that effective pandemic preparedness involves not only thorough planning but also decisive actions during implementation. Studies highlight the importance of fostering societal participation and consensus, executing timely and effective crisis communication, coordinating multi-sectoral actions, and maintaining rapid response capabilities. These dynamics support the importance of a proactive approach to global public health threats, emphasizing that mitigation and elimination strategies may yield varied results depending on each country's unique context. Thus, intervention frameworks must be adapted to align with each country's specific needs and capacities.

In conclusion, this study highlights substantial gaps in international preparedness for global health emergencies and advocates for a proactive management approach rather than reactive measures. By examining diverse country experiences, this study offers valuable insights for policymakers, guiding the development of more resilient and effective strategies for future health crises.

Author Contributions

All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by Faruk Yılmaz, İlhan Kerem Şenel, and Özgür İnce. The first draft of the manuscript was written by Faruk Yılmaz, and all authors commented on previous versions of the manuscript. Faruk Yılmaz and İlhan Kerem Şenel were involved in conceptualization. Faruk Yılmaz and Özgür İnce contributed to methodology. Faruk Yılmaz, İlhan Kerem Şenel, and Özgür İnce performed formal analysis and investigation. Faruk Yılmaz prepared original draft. Faruk Yılmaz, İlhan Kerem Şenel, and Özgür İnce were involved in review and editing. İlhan Kerem Şenel supervised the study. The corresponding author, Faruk Yılmaz, was actively involved at every stage of the research process, ensuring the integrity and accuracy of the work. All authors read and approved the final manuscript.

Acknowledgments

This article is derived from Faruk Yilmaz's 2023 Ph.D. dissertation titled "Assessing Countries' Efficiency in the Fight Against COVID-19 by Data Envelopment Analysis and Machine Learning," completed at the Department of Health Management, Institute of Graduate Studies, Istanbul University-Cerrahpaşa.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Aktas, E., F. Ülengin, I. Topcu, and E. H. Gundes. 2022. "The Efficiency of Nations in the Struggle Against the COVID-19 Pandemic." In *Handbook of Research on Healthcare Standards, Policies, and Reform*, 282–319. IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-8868-0.ch017.

Anderson, R. M., H. Heesterbeek, D. Klinkenberg, and T. D. Hollingsworth. 2020. "How Will Country-Based Mitigation Measures Influence the Course of the COVID-19 Epidemic?" *Lancet* 395, no. 10228: 931–934. https://doi.org/10.1016/S0140-6736(20)30567-5.

Aydin, N., and G. Yurdakul. 2020. "Assessing Countries' Performances Against COVID-19 via WSIDEA and Machine Learning Algorithms." *Applied Soft Computing* 97: 106792. https://doi.org/10.1016/j.asoc.2020.106792.

Banker, R. D., A. Charnes, and W. W. Cooper. 1984. "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science* 30, no. 9: 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45, no. 1: 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen. 1984. *Classification and Regression Trees*. Wadsworth.

Breitenbach, M. C., V. Ngobeni, and G. C. Aye. 2021. "Global Healthcare Resource Efficiency in the Management of COVID-19 Death and Infection Prevalence Rates." *Frontiers in Public Health* 9: 638481. https://doi.org/10.3389/fpubh.2021.638481.

Cameron, E. E., J. R. Nuzzo, and J. A. Bell. 2019. *GHS Index Building: Global Health Security Index, Collective Action and Accountability*. Johns Hopkins University.

Carrillo-Larco, R. M., and M. Castillo-Cara. 2020. "Using Country-Level Variables to Classify Countries According to the Number of Confirmed COVID-19 Cases: An Unsupervised Machine Learning Approach." *Wellcome Open Research* 5: 56. https://doi.org/10.12688/wellcomeopenres.15819.3.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research* 2, no. 6: 429–444. https://doi.org/10.1016/0377-2217(78)90138-8.

Cheshmehzangi, A., M. Sedrez, J. Ren, et al. 2021. "The Effect of Mobility on the Spread of COVID-19 in Light of Regional Differences in the European Union." *Sustainability* 13, no. 10: 5395. https://doi.org/10.3390/su13105395.

Delis, M. D., M. Iosifidi, and M. Tasiou. 2021. "Efficiency of Government Policy During the COVID-19 Pandemic." MPRA 107292, Munich Personal RePEc Archive. https://mpra.ub.uni-muenchen.de/107292/.

Delis, M. D., M. Iosifidi, and M. Tasiou. 2023. "Efficiency of Government Policy During the COVID-19 Pandemic." *Annals of Operations Research* 328, no. 2: 1287–1312. https://doi.org/10.1007/s10479-023-05364-9.

Ferguson, N. M., D. Laydon, G. Nedjati-Gilani, et al. 2020. "Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand." *Imperial College COVID-19 Response Team* 10: 77482.

Fernandez-Delgado, M., E. Cernadas, S. Barro, and D. Amorim. 2014. "Do We Need Hundreds of Classifiers to Solve Real World Classification Problems?" *Journal of Machine Learning Research* 15: 3133–3181.

Ghasemi, A., Y. Boroumand, and M. Shirazi. 2020. "How Do Governments Perform in Facing COVID-19?" *MPRA*: 99844.

Hale, T., N. Angrist, R. Goldszmidt, et al. 2021. "A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker)." *Nature Human Behaviour* 5, no. 4: 529–538. https://doi.org/10.1038/s41562-021-01079-8.

Hartshorn, S. 2016. *Machine Learning With Random Forests and Decision Trees: A Visual Guide For Beginners*. Amazon.

Hjerpe, A. 2016. *Computing Random Forests Variable Importance Measures (VIM) on Mixed Numerical and Categorical Data*. School of Computer Science and Communication (CSC).

Hua, Z., and Y. Bian. 2007. "DEA With Undesirable Factors." In *Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis*, edited by J. Zhu and W. D. Cook, 103–121. Springer US.

Ibrahim, M. D., F. A. Binofai, and R. MM Alshamsi. 2020. "Pandemic Response Management Framework Based on Efficiency of COVID-19 Control and Treatment." *Future Virology* 15, no. 12: 801–816. https://doi.org/10.2217/fvl-2020-0368.

Imtyaz, A., H. Abid Haleem, and M. Javaid. 2020. "Analysing Governmental Response to the COVID-19 Pandemic." *Journal of Oral Biology and Craniofacial Research* 10, no. 4: 504–513. https://doi.org/10.1016/j.jobcr.2020.08.005.

Jahanshahloo, G. R., F. H. Lotfi, N. Shoja, G. Tohidi, and S. Razavyan. 2005. "Undesirable Inputs and Outputs in DEA Models." *Applied Mathematics and Computation* 169, no. 2: 917–925. https://doi.org/10.1016/j.amc.2004.09.069.

Kelleher, J. D., B. Mac Namee, and A. D'Arcy. 2020. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. 2nd ed. The MIT Press.

Kohad, P. 2021. "K-Means clustering and Its Real World Use Case." https://www.linkedin.com/pulse/k-means-clustering-its-real-world-use-case-pratik-kohad-1c.

Legido-Quigley, H., N. Asgari, Y. Y. Teo, et al. 2020. "Are High-Performing Health Systems Resilient Against the COVID-19 Epidemic?" *Lancet* 395, no. 10227: 848–850. https://doi.org/10.1016/S0140-6736(20)30551-1.

Lotfi, M., M. R. Hamblin, and N. Rezaei. 2020. "COVID-19: Transmission, Prevention, and Potential Therapeutic Opportunities." *Clinica Chimica Acta* 508: 254–266. https://doi.org/10.1016/j.cca.2020.05.044.

Lupu, D., and R. Tiganasu. 2022. "COVID-19 and the Efficiency of Health Systems in Europe." *Health Economics Review* 12, no. 1: 14. https://doi.org/10.1186/s13561-022-00358-y.

Macqueen, J. 1967. "Some Methods for Classification and Analysis of Multivariate Observations." In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press.

Maimon, O. Z., and L. Rokach. 2014. *Data Mining With Decision Trees: Theory and Applications*. 2nd ed. World Scientific.

Martínez-Córdoba, P. J., B. Benito, and I. M. García-Sánchez. 2021. "Efficiency in the Governance of the Covid-19 Pandemic: Political and Territorial Factors." *Globalization and Health* 17, no. 1: 113. https://doi.org/10.1186/s12992-021-00759-4.

Ordu, M., H. Kirli Akin, and E. Demir. 2021. "Healthcare Systems and Covid19: Lessons to Be Learnt From Efficient Countries." *International Journal of Health Planning and Management* 36, no. 5: 1476–1485. https://doi.org/10.1002/hpm.3187.

Ourworld. (2022), "Coronavirus Pandemic (COVID-19), Our World Data 2020." https://ourworldindata.org/coronavirus.

Our World in Data. 2023. "COVID-19 Data." *GitHub*. https://github.com/owid/covid-19-data/tree/master/public/data.

Ozcan, Y. A. 2014. "Advanced DEA Models." In *Health Care Benchmarking and Performance Evaluation: An Assessment Using Data Envelopment Analysis (DEA)*, edited by Y. A. Ozcan, 121–137. Springer US.

Pan, A., L. Liu, C. Wang, et al. 2020. "Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China." *Journal of the American Medical Association* 323, no. 19: 1915–1923. https://doi.org/10.1001/jama.2020.6130.

Pereira, M. A., D. C. Dinis, D. C. Ferreira, J. R. Figueira, and R. C. Marques. 2022. "A Network Data Envelopment Analysis to Estimate Nations' Efficiency in the Fight Against SARS-CoV-2." *Expert Systems With Applications* 210: 118362. https://doi.org/10.1016/j.eswa.2022.118362.

Ranney, M. L., V. Griffeth, and A. K. Jha. 2020. "Critical Supply Shortages—The Need for Ventilators and Personal Protective Equipment During the Covid-19 Pandemic." *New England Journal of Medicine* 382, no. 18: e41. https://doi.org/10.1056/NEJMp2006141.

Rizvi, R. F., K. J. T. Craig, R. Hekmat, et al. 2021. "Effectiveness of Non-Pharmaceutical Interventions Related to Social Distancing on Respiratory Viral Infectious Disease Outcomes: A Rapid Evidence-Based Review and Meta-Analysis." *SAGE Open Medicine* 9: 1–13. https://doi.org/10.1177/20503121211022973.

Rodriguez, M. Z., C. H. Comin, D. Casanova, et al. 2019. "Clustering Algorithms: A Comparative Approach." *PLoS One* 14, no. 1: e0210236. https://doi.org/10.1371/journal.pone.0210236.

Rokach, L. 2016. "Decision Forest: Twenty Years of Research." *Information Fusion* 27: 111–125. https://doi.org/10.1016/j.inffus.2015.06.005.

Seiford, L. M., and J. Zhu. 2002. "Modeling Undesirable Factors in Efficiency Evaluation." *European Journal of Operational Research* 142, no. 1: 16–20. https://doi.org/10.1016/S0377-2217(01)00293-4.

Singh, S., and M. Giri. 2014. "Comparative Study Id3, Cart and C4.5 Decision Tree Algorithm: A Survey." *International Journal of Advanced Information Science and Technology* 3, no. 7: 47–52. https://doi.org/10.15693/ijaist/2014.v3i7.

Su, E. C. Y., C. H. Hsiao, Y. T. Chen, and S. H. Yu. 2021. "An Examination of COVID-19 Mitigation Efficiency Among 23 Countries." *Healthcare* 9, no. 6: 755. https://doi.org/10.3390/healthcare9060755.

Sullivan, W. 2017. *Machine Learning Beginners Guide Algorithms: Supervised & Unsupervised Learning, Decision Tree & Random Forest Introduction*. Healthy Pragmatic Solutions Inc.

Taherinezhad, A., and A. Alinezhad. 2022. "COVID-19 Crisis Management: Global Appraisal Using Two-Stage DEA and Ensemble Learning Algorithms." *Scientia Iranica*. In press. https://doi.org/10.24200/sci.2022.58911.5962.

United Nations: Human Development Index (HDI). 2020. "Human Development Index." February 8, 2023. https://hdr.undp.org/data-center/human-development-index#/indicies/HDI.

Walker, P. G. T., C. Whittaker, O. J. Watson, et al. 2020. "The Impact of COVID-19 and Strategies for Mitigation and Suppression in Low- and Middle-Income Countries." *Science* 369, no. 6502: 413–422. https://doi.org/10.1126/science.abc0035.

World Bank. 2023. "World Bank Open Data." https://data.worldbank.org.

World Health Organization. 2016. *Health Workforce Requirements for Universal Health Coverage and the Sustainable Development Goals*. Human Resources for Health Observer, 17, World Health Organization.

World Health Organization. 2020. "WHO Director-General's Opening Remarks at the Media Briefing on COVID-19—11 March 2020." https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

World Health Organization. 2023. "Global Health Expenditure Database." https://apps.who.int/nha/database.

Xu, Y., Y. S. Park, and J. D. Park. 2021. "Measuring the Response Performance of U.S. States Against COVID-19 Using an Integrated DEA, CART, and Logistic Regression Approach." *Healthcare* 9, no. 3: 268. https://doi.org/10.3390/healthcare9030268.

Yilmaz, F. 2023. Unpublished doctoral dissertation. "Assessing Countries' Efficiency in Fight Against COVID-19 by Data Envelopment Analysis and Machine Learning." İstanbul University-Cerrahpaşa.

You, S., and H. Yan. 2011. "A New Approach in Modelling Undesirable Output in DEA Model." *Journal of the Operational Research Society* 62: 2146–2156. https://doi.org/10.1057/jors.2011.1.

Zhou, Z. H. 2021. *Machine Learning*. Springer Nature.

Zhu, Q., X. Zhou, and S. Liu. 2022. "A Multi-Stage Super DEA Efficiency Evaluation Model of COVID-19 Pandemic Transmission Performance." *WHICEB 2022 Proceedings*, 73.