

NRC Publications Archive Archives des publications du CNRC

Longitudinal bi-criteria framework for assessing national healthcare responses to pandemic outbreaks

Guitouni, Adel; Belacel, Nabil; Benabbou, Loubna; Moa, Belaid; Erman, Munire; Abdul, Halim

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1038/s41598-024-69212-x>

Scientific Reports, 14, 1, 2024-09-27

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

<https://nrc-publications.canada.ca/eng/view/object/?id=bc01531f-1cea-443d-a1e6-63edc3625a1c>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=bc01531f-1cea-443d-a1e6-63edc3625a1c>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



OPEN

Longitudinal bi-criteria framework for assessing national healthcare responses to pandemic outbreaks

Adel Guitouni^{1✉}, Nabil Belacel^{2✉}, Loubna Benabbou³, Belaid Moa⁴, Munire Erman⁵ & Halim Abdul⁶

Pandemics like COVID-19 have illuminated the significant disparities in the performance of national healthcare systems (NHCSs) during rapidly evolving crises. The challenge of comparing NHCS performance has been a difficult topic in the literature. To address this gap, our study introduces a bi-criteria longitudinal algorithm that merges fuzzy clustering with Data Envelopment Analysis (DEA). This new approach provides a comprehensive and dynamic assessment of NHCS performance and efficiency during the early phase of the pandemic. By categorizing each NHCS as an efficient performer, inefficient performer, efficient underperformer, or inefficient underperformer, our analysis vividly represents performance dynamics, clearly identifying the top and bottom performers within each cluster of countries. Our methodology offers valuable insights for performance evaluation and benchmarking, with significant implications for enhancing pandemic response strategies. The study's findings are discussed from theoretical and practical perspectives, offering guidance for future health system assessments and policy-making.

Keywords Bi-criteria analysis, Healthcare performance evaluation, Data envelopment analysis, Fuzzy clustering, Longitudinal analysis, COVID-19 pandemic

Which country's national healthcare system (NHCS) responded better to the outbreak of the COVID-19 pandemic? The novel coronavirus 2019 (COVID-19/SARS-CoV2) outbreak, often called the pandemic of the century, has profoundly impacted healthcare systems worldwide since it was first identified in Wuhan, China^{1,2}. Characterized by high transmission and fatality rates, COVID-19 has disrupted the normal functioning of healthcare systems globally³.

At the peak of each wave, public health authorities cancelled all non-urgent healthcare services to free up capacity for the expected surge in COVID-19 patients, aiming to prevent healthcare system collapse, as witnessed in countries like Italy⁴ and India⁵. Countries implemented various measures to suppress the spread of COVID-19, maintain healthcare capacity, deliver high-quality care, and reduce secondary mortality rates. However, physical distancing, movement restrictions, and the cancellation of non-urgent services led to other healthcare problems, such as reduced access to cancer treatments, increased anxiety, mental exhaustion, and backlogged waitlists^{6–8}.

Significant differences in responses and outcomes across countries with varying death and recovery rates underscore the need for systematic approaches to assess and benchmark NHCS performance. For instance, France and Germany exhibited vastly different death and recovery rates during the pandemic's first 100 days despite similar geographic and socio-economic conditions⁹. Similarly, Canada and the USA showed significant NHCS performance and efficiency variations despite substantial resource allocation (see Fig. 1). This has motivated researchers to develop methodological approaches to compare NHCS performance during global pandemics like COVID-19.

International comparisons of NHCSs improve health services, enforce accountability, share strategies, and encourage mutual learning¹⁰. However, such comparisons face significant challenges, including defining the

¹Gustavson School of Business, University of Victoria, Victoria, BC, Canada. ²Digital Technologies Research Center, National Research Council, Ottawa, ON, Canada. ³Department of Management Sciences, Université du Québec à Rimouski, Rimouski, QC, Canada. ⁴Advanced Computing, University of Victoria, Victoria, BC, Canada. ⁵Respiratory Therapy, Medical Day Care, Cancer Care, Social Work, Maternity and Pediatrics Units, Campbell River General Hospital, Campbell River, BC, Canada. ⁶Department of Economics, University of Victoria, Victoria, BC, Canada. ✉email: adelg@uvic.ca; nabil.belacel@nrc-cnrc.gc.ca

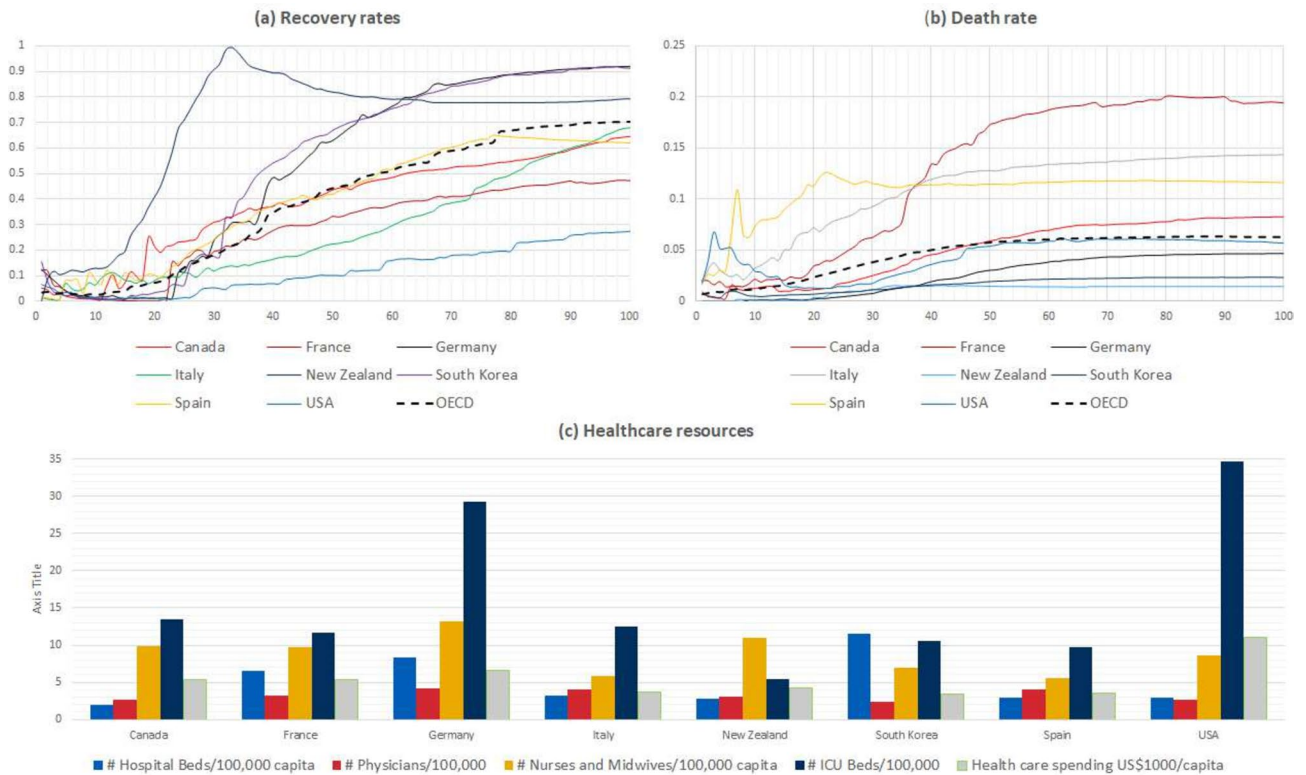


Figure 1. Recovery and death rates, and healthcare resources from selected OECD countries: Canada, France, Germany, Italy, New Zealand, South Korea, and the USA: (a) daily cumulative recovery rates, (b) daily cumulative death rates, (c) resources of NHCS.

boundaries of the health system, selecting performance metrics, ensuring comparability across borders, and accounting for dynamic responses over time. The complexity is compounded by the tension between assessing the entire system versus fragmented comparisons of health units (e.g., hospitals). Various international organizations, such as the World Health Organization, the Organization for Economic Co-operation and Development (OECD), and the European Union, have adopted system assessments to influence public policies and accelerate debates about NHCS performance¹⁰.

Interest in comparing international health systems is growing as NHCSs face increasing pressure to improve performance^{10,11}. However, defining the unit of analysis is challenging. According to the World Health Organization (WHO)¹², a health system encompasses all organizations, agents, and actions involved in healthcare provision, public health promotion, and protection. Defining national health system boundaries is complex due to varying perspectives and levels of analysis^{13,14}. Additionally, the objectives and terminology used to describe healthcare systems are often ambiguous^{15,16}.

Comparing NHCS performance during a global pandemic is a complex and challenging task. The pandemic simultaneously impacts health, social, and economic systems, making isolating healthcare system effects difficult. Moreover, there needs to be standardized criteria or performance metrics for comparing NHCS performances, adding to the complexity. The dynamic nature of NHCS responses, which evolve as systems learn and adapt, further complicates comparisons. This complexity presents an intriguing challenge for researchers, policymakers, and professionals in health systems research, health policy analysis, and health economics.

Comparing health systems' performance is a complex task studied extensively in fields such as health systems research, health policy analysis, and health economics^{13,17,18}. Additionally, operational research techniques address issues like health system bottlenecks and low hospital performance¹⁹. These techniques have also been pivotal in solving hospital and health system challenges related to COVID-19^{20–24}. However, unlike these recent works, our proposed approach first combines fuzzy clustering to group countries into homogeneous and distinct clusters, taking into account their population health similarities. Then, we propose a targeted application of longitudinal Data Envelopment Analysis (DEA)²⁵ to comparable countries within each cluster. Finally, we propose a qualitative bi-criteria analysis to capture the dynamic the performance of each country's response over time considering efficiency and clinical performances.

This study introduces an innovative bi-criteria method to dynamically compare NHCS responses to the COVID-19 outbreak. The proposed method, which combines fuzzy clustering and longitudinal data envelopment analysis (DEA), provides a comprehensive and dynamic assessment of NHCS performance and efficiency during the outbreak phase. Clinical outcomes are the first criterion for assessing NHCS performance. At the same time, longitudinal DEA evaluates response efficiency with five inputs representing the resources available to the NHCS and two outputs representing the ratios of recovered and the number of survived cases to the total number of infected individuals in a population. The algorithm categorizes each NHCS as an efficient-performer,

inefficient-performer, efficient-under-performer, or inefficient-under-performer for the pandemic's first 100 days. The study visually represents performance dynamics over time, highlighting each cluster's best and worst performers during the early outbreak phase. This innovative algorithm has potential applications in various domains and levels of analysis, particularly in complex performance evaluation and benchmarking. The implications for performance assessment and benchmarking are discussed theoretically and practically, demonstrating the practicality and innovation of this approach for researchers, policymakers, and professionals in health systems research, health policy analysis, and health economics.

This paper is structured as follows: **Related work** section provides a review of the related literature. **Methodology: qualitative bi-criteria approach for the comparison of NHCS** section, the methodology, develops the bi-criteria approach for comparing NHCSs. **Results** section presents the empirical results for the first 100 days of the pandemic. **Discussion and managerial implications** section discusses the theoretical and practical implications of the work. Finally, **conclusion** in section presents the conclusions, summarizing the essential findings and their implications.

Related work

Comparing the performance of health systems is a persistent challenge in several streams of literature^{13,17,18}. With the COVID-19 pandemic, the need for systematic approaches to assess and benchmark national health systems (NHCS) performance has become even more pressing¹⁰. The global nature of the pandemic provides a unique context for comparing NHCS responses facing almost the same disruption simultaneously.

Our study compares the responses of NHCSs from 37 member countries of the Organization for Economic Co-operation and Development (OECD), driven by the availability of reliable data. We acknowledge that data from other countries may be less readily available or trustworthy, and contextual differences can introduce further complications. For example, populations in different countries may be vulnerable to COVID-19, with front-line healthcare workers at a higher risk of infection²⁶. Additionally, demographic groups such as women have lower infection and mortality rates than men²⁷. The mean age of death due to COVID-19 is around 82 years^{28–30}, and pre-existing comorbidities are associated with higher mortality rates^{31–34}. Furthermore, populations with higher vulnerability to COVID-19 are reported to require more intensive care²⁹ and experience higher mortality rates²⁸.

Our work builds on existing literature evaluating NHCS, which typically considers three main performance criteria: access to healthcare services, quality of care, and efficiency (cost)^{35,36}. However, there needs to be more clarity regarding the definition of these domains, how they overlap, and how they relate to critical health system objectives^{15,16,37}. Additionally, the distribution of performance objectives across different population groups adds complexity to evaluating healthcare performance³⁸.

Two primary methodological approaches have been used to assess NHCS efficiency. The first is the ratio analysis approach, based on individual partial indicators often developed as ratios to provide different levels of efficiency^{39,40}. This approach can lead to conflicting results due to the need to calculate multiple ratios to capture various performance dimensions⁴¹. The second approach uses non-parametric econometric techniques like DEA, which computes a comprehensive efficiency index^{42–45}. Our work is grounded in this second stream, as DEA has been widely used to assess the relative performance of healthcare organizations and identify ways to improve their efficiency⁴⁶. For example, DEA has been applied to estimate the technical efficiency of hospitals in Turkey⁴⁷, assess healthcare center performance¹⁷, evaluate health insurance reform in China⁴⁸, and measure the efficiency of rural hospitals in Missouri⁴⁹. DEA has also been used to assess the performance of healthcare systems during the COVID-19 pandemic^{19,50–55}.

Comparing NHCS responses to COVID-19 is a multi-criteria problem, requiring consideration of multiple performance dimensions. Multiple criteria decision analysis (MCDA) addresses complex decision-making problems involving conflicting and non-commensurable evaluations^{56–58}. Most MCDA methods construct representations in the preference space and use different preference aggregation procedures to synthesize partial preference structures into an aggregated preference structure^{59,60}. However, the comparison of NHCS does not fit within traditional MCDA literature^{56,57,61} for several reasons: there is no single decision-maker guiding the analysis, the output is not a traditional decision problem, and it is crucial to capture the dynamics of NHCS performance over time.

No existing method combines longitudinal DEA and fuzzy clustering to create a qualitative, bi-criteria decision analysis framework for comparing NHCS responses to the COVID-19 pandemic. Our approach addresses this gap by providing a comprehensive and dynamic NHCS performance and efficiency assessment.

Methodology: qualitative bi-criteria approach for the comparison of NHCS

We present a novel visual and qualitative bi-criteria method for comparing the dynamic responses of national healthcare systems (NHCSs) to the COVID-19 pandemic. To address the challenges inherent in such comparisons, we propose a conceptual model of an NHCS, illustrated in Fig. 2. This model encapsulates the fundamental structural components of the behavioural and ecological model of health services, positioning the NHCS as a core element of the national health system. It comprises four interdependent, integrated, and dynamic components: (1) governance of the national health system, (2) public health, (3) NHCS, and (4) outcome management (accountability).

The model posits that national governance, encompassing public health strategy, ecosystem, leadership, partnerships, investment, supply chain, and information management, defines public health policies, programs, and actions while managing funding, risk, and accountability⁶². Public health then formulates, adapts, and enforces policies to contain the virus, improve population immunization, and enhance testing and tracing efforts. Concurrently, national governance directs the NHCS's preparedness and response to the pandemic.

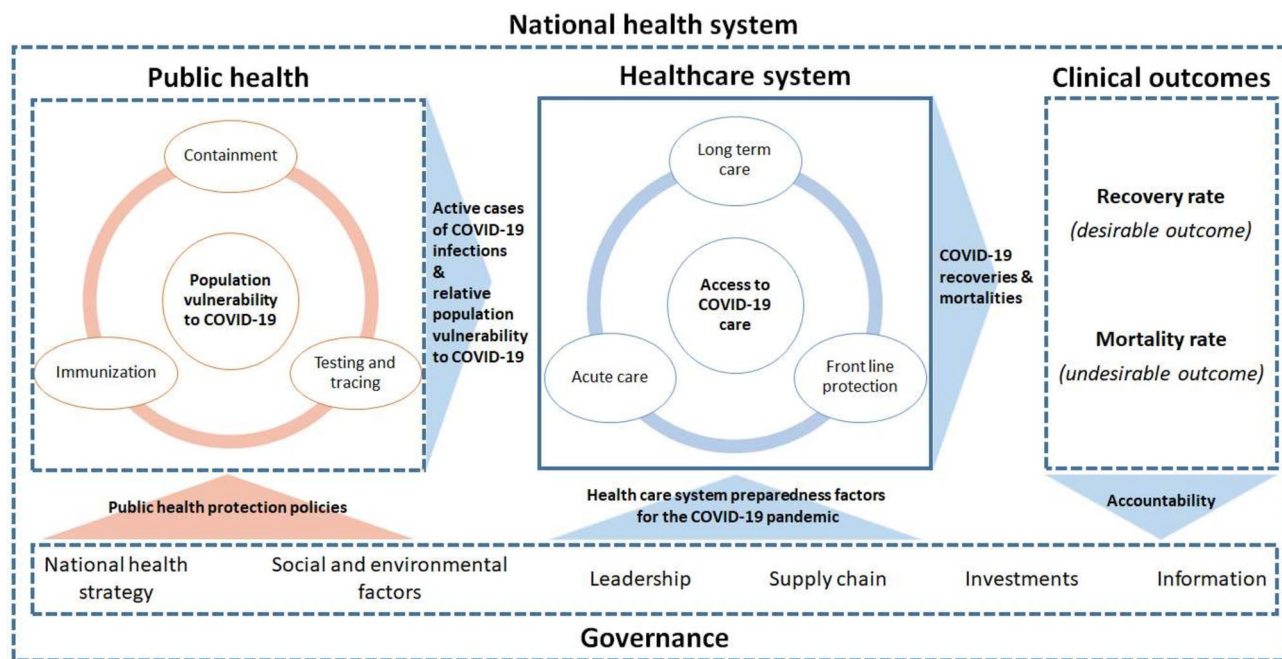


Figure 2. Conceptual model of an NHCS in response to the COVID-19 pandemic.

The model further suggests that the complex interactions between social and environmental factors—such as predispositions, behaviours, and needs—and the virus’s spread influence the severity of cases managed by the NHCS. It builds on the premise that various public health factors moderate the impact of the virus on population health and access to healthcare resources⁶³. In response to the COVID-19 outbreak, we conceptualize an NHCS as a collection of institutions tasked with restoring and preserving health. The primary capabilities of an NHCS in responding to COVID-19 include providing access to acute and long-term care and protecting frontline workers.

Our visual and qualitative bi-criteria method for comparing NHCS responses to the COVID-19 pandemic is specifically designed to address the unique challenges of such comparisons, including the lack of homogeneity among NHCS units, the multidimensional nature of performance metrics, and the need to capture performance dynamics over time. The NHCS is treated as a decision-making unit (DMU) in this analysis. The Algorithm 1 integrates longitudinal DEA and fuzzy clustering to develop an objective and interpretable framework.

-
- 1: **Input:**
 - 2: $DMU = \{DMU_1, \dots, DMU_n\}$ set of n NHCS represented by longitudinal input dimensions $I(DMU_i, t)$ and output dimensions $O(DMU_i, t)$
 - 3: **Run Longitudinal DEA**
 - 4: **while** $t \leq T$ and $i \leq n$ **do**
 - 5: Calculate $DEA(DMU_i, t)$, $i \leftarrow i + 1$, $t \leftarrow t + 1$
 - 6: **Run the fuzzy clustering algorithm FJM**
 - 7: **while** $i \leq n$ **do**
 - 8: Determine (DMU_i, V_i) , $i \leftarrow i + 1$
 - 9: **Bi-criterion comparison of DMU**
 - 10: Calculate $Efficiency(DMU_i, t)$
 - 11: Calculate $Performance(DMU_i, t)$
 - 12: **if** $(DMU_i, DMU_j) \in V^2$ **then**
 - 13: Compare DMU_i to DMU_j
-

Algorithm 1. Qualitative bi-criteria approach.

Longitudinal data envelopment analysis

We use DEA to compare the efficiency of DMUs. DEA uses linear programming to create an efficiency frontier by optimizing the weighted output-to-input ratio of each DMU⁶⁴.

Definition of inputs: NHCS preparedness and COVID-19 cases

Resources available to NHCS are the inputs. Resource shortages limit the ability of an NHCS to respond to the pandemic^{65,66}. The global nature of the crisis means that “we are all in it together”, and access to multilateral support and assistance is unlikely. Therefore, we assume that resource preparedness of an NHCS for the pandemic is a critical determinant of its performance. The recent literature on COVID-19 identifies five resource dimensions (X_1 to X_5) of NHCS preparedness for the COVID-19 pandemic (see Table 1), which are X_1 to X_5 . The other input considered in this study is the longitudinal data on the cumulative number of confirmed cases per day for each

Notation	Preparedness dimension for all countries	Reference(s)
X_1 :	Number of hospital beds per capita	⁶⁸
X_2 :	Number of physicians per capita	⁶⁹
X_3 :	Number of nurses and midwives per capita	⁶⁹
X_4 :	Number of intensive care beds per capita	⁷⁰
X_5 :	Health expenditures (USD) per capita	⁷¹

Table 1. Input dimensions for longitudinal data envelopment analysis.

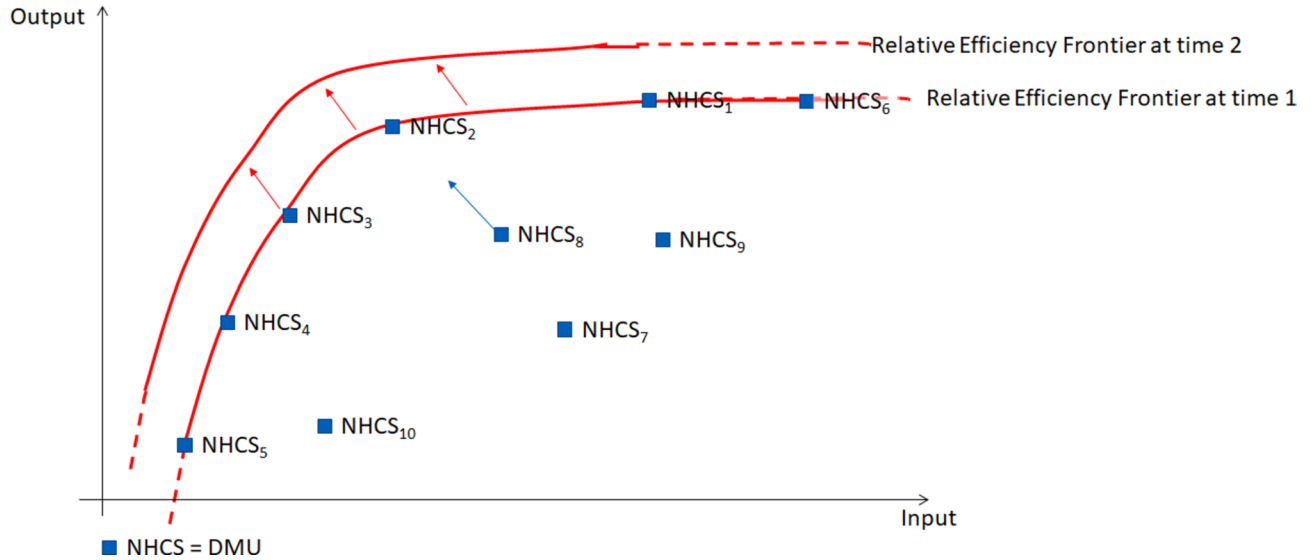


Figure 3. Principle of longitudinal DEA.

country (X_6^t). A confirmed COVID-19 case is a person with laboratory confirmation of COVID-19 infection based on a validated and approved test performed in the community at a hospital or reference laboratory⁶⁷.

Definition of outputs: NHCS clinical outcomes

We consider two outputs: survival and recovery rates. The first longitudinal output $y_1(DMU_i, t)$ is the infection recovery rate (IRR) at time t , which indicates what percentage of total diagnosed positive cases are cured of the disease (see Eq. 1).

$$y_1(DMU_i, t) = IRR(DMU_i, t) = \frac{(\text{Total number of recovered cases})_i^t}{(\text{Total number of confirmed cases})_i^t} \tag{1}$$

for a country i at time t . The $IRR(DMU_i, t)$ is an estimated metric because no country has succeeded in testing the entirety of its population.

We use the case fatality rate (CFR) at time t to estimate the death rate (see Eq. 2).

$$CFR(DMU_i, t) = \frac{(\text{Total number of deaths})_i^t}{(\text{Total number of confirmed cases})_i^t} \tag{2}$$

for a country i at time t . We define the survival rate as the second longitudinal output $y_2(DMU_i, t)$ (see Eq. 3).

$$y_2(DMU_i, t) = 1 - CFR(DMU_i, t) \tag{3}$$

It is worth emphasizing that CFR does not involve a lag, as we are interested in the instantaneous response of an NHCS.

Longitudinal DEA

We focus on the response of the healthcare system to the outbreak of COVID-19. The inputs are defined by the preparedness dimensions and COVID-19-related demand for healthcare services, while the outputs are specified by the clinical outcomes. Given that the outputs do not necessarily change in a proportional way to the inputs, we assume variable returns to scale (VRS) technology^{43,72}. However, the inputs and outputs vary over time during the pandemic. Therefore, we expect the efficiency of an NHCS to change over time. The longitudinal DEA is,

therefore, introduced to capture these changes (as shown in Fig. 3). The efficiency frontier at time $t = 1$ is defined by NHCSs 1 to 6. At time $t = 2$, the efficiency frontier may change as NHCSs improve their institutional or technical performance, which explains the efficiency frontier shift.

Let (X^t, Y^t) be two longitudinal matrices representing inputs and outputs, respectively. As the study's objective revolves around the NHCS's performance in using its resources, an input-oriented VRS model⁷³ is adopted to measure the efficiency of each DMU_i for $i = 1, \dots, 37$ at time t . Given that we are dealing with ratio outputs, we adopt the model proposed by Emrouznejad et al.⁷⁴ to take into account the ratio nature of the two outputs. The model is given in Eq. 4, with $t = 1, \dots, T$. To preserve the discriminatory power of the DEA⁷⁵, as explained [Bi-criteria visual analysis of efficiency and performance of NHCS](#) in section, we make sure that the number of DMUs is higher than the numbers of inputs and outputs.

$$\begin{aligned} & \min_{\theta^t, \lambda^t} \{\theta^t\} \\ & \text{subject to} \\ & \theta^t x_0^t - X^t \lambda^t \geq 0 \\ & N^t \lambda^t - y_0^t D^t \lambda^t \geq 0 \\ & e^T \lambda^t = 1 \\ & \lambda^t \geq 0 \end{aligned} \quad (4)$$

With N^t and D^t are the longitudinal matrices of the numerator and denominator for the ratio output Y^t at time $t = 1, \dots, T$, i.e. for each $DMU_i, i = 1 \dots 37$:

$$y_1(DMU_i, t) = \frac{(\text{Total number of recovered cases})_i^t}{(\text{Total number of confirmed cases})_i^t} = \frac{n_1(DMU_i, t)}{d(DMU_i, t)} \quad (5)$$

and

$$y_2(DMU_i, t) = \frac{1 - (\text{Total number of deaths})_i^t}{(\text{Total number of confirmed cases})_i^t} = \frac{n_2(DMU_i, t)}{d(DMU_i, t)} \quad (6)$$

Various extensions have been proposed to overcome the limitations of the original standard DEA models. For instance, Aleskerov and Petrushchen⁷⁶ suggest that non-homogeneous DMUs may present challenges when interpreting results. One solution proposed in the literature is to use a combination of clustering and DEA to evaluate the efficiency and productivity of DMUs⁷⁷. For example, Dai and Kuosmanen⁷⁸ applied a combination of clustering and DEA to identify DMUs with similar characteristics, resulting in more efficient and meaningful results than traditional methods. Similarly, other studies have combined k-means clustering with DEA to evaluate public hospitals⁷⁹ and applied a fuzzy clustering cooperative game with DEA to assess hospital efficiency⁸⁰. Zarrin et al.⁸¹ used a combination of self-organizing map artificial neural network, cluster analysis, and slacks-based measure DEA with multilayer perceptron neural network to analyze the efficiency of over 1100 hospitals in Germany.

The assessment and comparison of NHCS performance over time is crucial in understanding how different countries are responding to the pandemic outbreak. While various metrics exist to assess the performance of DMUs as documented by⁸², the Malmquist index (M) offers distinct advantages. Unlike traditional metrics that might only compare a unit's efficiency to an average performer, the M index leverages the concept of a production frontier to enable a more rigorous benchmark. Additionally, some metrics solely capture efficiency changes, overlooking advancements due to technological progress. The Malmquist index addresses this shortcoming by incorporating both technical change and efficiency change, as highlighted by⁸³. Consequently, the Malmquist index provides a richer and more nuanced picture of performance change over time, making it a valuable tool for DEA researchers to dynamically evaluate DMUs. The conditions of the pandemic, combined with the adaptation to change, may result in performance gains or losses over time, depending on how NHCS adapt to various external and internal factors. To measure the change in productivity of each NHCS over time, we use the Malmquist productivity index M. This method compares the relative performance of an NHCS with respect to reference states from two different time periods. It also allows for decomposition of productivity change into institutional efficiency change and technical change⁸⁴, making it particularly useful for analyzing healthcare performance over time^{85,86}.

To analyze the change in performance between two periods, we denote the efficiency scores for each period as $\theta^{1,1}$ and $\theta^{2,1}$, where the first exponent refers to the period and the second to the state of technology. The MTEC, which measures the catch-up effect between two consecutive periods, is given by Eq. (7). The MTC, which measures the frontier-shift effect or the change in the reference frontier between two consecutive periods, is given

by Eq. (8). Using the period 2 observation as the reference benchmark to evaluate the shift in the frontier²⁵, the Malmquist productivity index (M) can be calculated using Eq. (9).

$$MTEC = \theta^{2,1} / \theta^{1,1} \quad (7)$$

$$MTC = \left[\left(\frac{\theta^{2,1}}{\theta^{2,2}} \right) * \left(\frac{\theta^{1,1}}{\theta^{1,2}} \right) \right]^{1/2} \quad (8)$$

$$M = MTEC * MTC = \left(\frac{\theta^{2,2}}{\theta^{1,1}} \right) * \left[\left(\frac{\theta^{2,1}}{\theta^{2,2}} \right) * \left(\frac{\theta^{1,1}}{\theta^{1,2}} \right) \right]^{1/2} \quad (9)$$

Finally, we utilize the Malmquist productivity index M to assess and compare the performance of NHCSs over time. This approach captures the dynamics of the response to the pandemic in different countries and allows for an analysis of performance gains or losses over time. Additionally, we decompose the long-term productivity change into consecutive sub-periods using the transitivity property of the Malmquist index. This enables us to study the temporal performance and efficiency evolution of NHCSs in response to COVID-19 outbreak through a combination of graphical analysis and the Malmquist method.

Fuzzy clustering: controlling for the population vulnerability

The vulnerability of a population to COVID-19 plays a significant role in determining the demand for national healthcare services. Countries with populations that are more vulnerable to the virus, such as those with a larger aging population, may require more resources to provide care and may experience higher mortality rates. To account for this, we use a clustering method to group countries with similar levels of population vulnerability to the pandemic. This allows us to control for the moderating effect of vulnerability on the demand for healthcare resources. The literature identifies 15 dimensions of vulnerability to COVID-19, which are detailed in Table 2.

The FJM approach is a fuzzy clustering method that is used in this study to NHCSs into homogeneous groups based on the vulnerability of their respective populations to COVID-19. FJM is a variation of the popular fuzzy clustering method, Fuzzy C-Means, which addresses the limitations of traditional clustering methods by allowing for complex relationships between entities and reducing data noise. FJM uses all possible centroid-to-pattern relocations in order to construct move-defined neighborhoods. This allows for a more efficient optimization of the membership degrees and centroids, resulting in better solutions compared to traditional hard clustering methods such as k-means⁹⁶. The FJM method has been proven to provide better results in multiple machine learning and pattern recognition applications^{97–99}. The FJM algorithm used in this study is summarized in Algorithm 2^{96,97} and the python code is available on GitHub at <https://github.com/nbelacel/FuzzyJMeansVNS>. The output of the FJM method is a matrix of membership degrees that describes the level of similarity between each NHCS and each cluster centroid, allowing for the assignment of NHCSs to multiple clusters simultaneously with different membership degrees.

Notation	Dimension of vulnerability	Supporting reference(s)
V ₁ :	Asthma	87
V ₂ :	Chronic obstructive pulmonary diseases	88,89
V ₃ :	Diabetes	32,35
V ₄ :	Cardiac diseases in male population	34,90
V ₅ :	Cardiac diseases in female population	31,34
V ₆ :	Peripheral vascular disease	34,91
V ₇ :	Hypertension in male population	32,33
V ₈ :	Hypertension in female population	32,33
V ₉ :	Population aged more than 65 years	28,29
V ₁₀ :	Smokers	89
V ₁₁ :	Obesity in male population	92,93
V ₁₂ :	Obesity prevalence in female population	92,93
V ₁₃ :	Cancer and immunosuppressive diseases	32
V ₁₄ :	Chronic kidney disease in population	32
V ₁₅ :	Density of the urban population	94,95

Table 2. Dimension of vulnerability of the population to the COVID-19.

1: **Input:**
2: $X = \{x_1, \dots, x_n\}$ set of n entities where x_i is a d -dimensional point
3: K : the number of clusters given a priori
4: **Output:**
5: $V = \{v_1, v_2, \dots, v_k\}$ set of k centroids and $U = [u_{ij}]_{n \times k}$ membership degree of each entity to each cluster
6: **Step 0: Initialization:**
7: Find initial centroids V
8: Fix value for $m \leftarrow 2$ /* m is the fuzziness parameter*/
9: Calculate the objective function R and $R_{opt} = R$
10: Fix value for $MaxIt$ /*Maximum of iterations*/ and $t \leftarrow 0$
11: **Step 1: Drop least useful centroid**
12: Find one centroid $v^{del} \in V$ to be deleted using some criteria
13: **Step 2: Unoccupied entities**
14: Find unoccupied entities /*entities that do not coincide with a cluster centroid*/
15: **Step 3: Add most useful entity to deleted centroid**
16: Select among the unoccupied entities, the one that will be added v^{ins} according to some criterion
17: $V \leftarrow V \cup \{v^{ins}\} \setminus \{v^{del}\}$
18: **Step 4: Local improvements**
19: Calculate Fuzzy membership U with new centroid set V calculated in Step 3.
20: Calculate the objective function R
21: **Step 5: Check for improvement in R**
22: **if** ($R < R_{opt}$) **AND** ($it \leq MaxIt$) **then**
23: Save the new currently best solution and its value:
24: $V_{opt} \leftarrow V; R_{opt} \leftarrow R$
25: $it \leftarrow it + 1$ and **Goto** Step2
26: **else**
27: Calculate the membership matrix U_{opt}
28: Stop with a local minimum (V_{opt}, U_{opt})
29: **End**

Algorithm 2. The fuzzy clustering algorithm FJM.

Bi-criteria visual analysis of efficiency and performance of NHCS

The performance of an NHCS in response to COVID-19 outbreak is conceptualized based on the clinical outcomes: maximizing the COVID-19 recovery rate and minimizing the COVID-19 death rate. The efficiency is measured using the DEA method. Together, these two categories define the qualitative multiple criteria map shown in Fig. 4. Thus, four cases can be considered for the DMU_i . First, a DMU_i is considered an *efficient-performer* if it outperforms similar countries in terms of clinical outcomes and it is efficient regarding its DEA efficient score at time t . Secondly, DMU_i is considered an *efficient-under-performer*, if it underperforms similar countries in terms of clinical outcomes and it is efficient at time t . Third, a DMU_i is considered an *inefficient-performer*, if it outperforms similar countries in terms of clinical outcomes but it is inefficient at time t . Finally, a DMU_i is considered an *inefficient-under-performer*, if it underperforms similar countries in terms of clinical outcomes and it is inefficient at time t . We have used visualization techniques to compare the performance and efficiency of different NHCSs in response to COVID-19 outbreak. To further clarify, we have established the following qualitative categories of performance (see Figs. 10, 11, 12, 13):

$$\text{if } \begin{cases} IRR(DMU_i, t) \geq IRR(V_c, t) \\ \text{and} \\ CFR(DMU_i, t) \leq CFR(V_c, t) \end{cases} \Rightarrow \text{Performer } DMU_i \quad (10)$$

$$\text{if } \begin{cases} IRR(DMU_i, t) < IRR(V_c, t) \\ \text{and} \\ CFR(DMU_i, t) > CFR(V_c, t) \end{cases} \Rightarrow \text{Under-performer } DMU_i \quad (11)$$

$$\text{if } \begin{cases} IRR(DMU_i, t) < IRR(V_c, t) \\ \text{and} \\ CFR(DMU_i, t) \leq CFR(V_c, t) \end{cases} \Rightarrow \text{Quasi-performer } DMU_i \quad (12)$$

$$\text{if } \begin{cases} IRR(DMU_i, t) \geq IRR(V_c, t) \\ \text{and} \\ CFR(DMU_i, t) > CFR(V_c, t) \end{cases} \Rightarrow \text{Pseudo-performer } DMU_i \quad (13)$$

where $IRR(V_c, t)$ and $CFR(V_c, t)$ are the average recovery and death rates at time t for the respective cluster V_c of similar countries. Similarly, we define the efficiency of DMU_i at time t based on DEA outputs as follows²⁵:

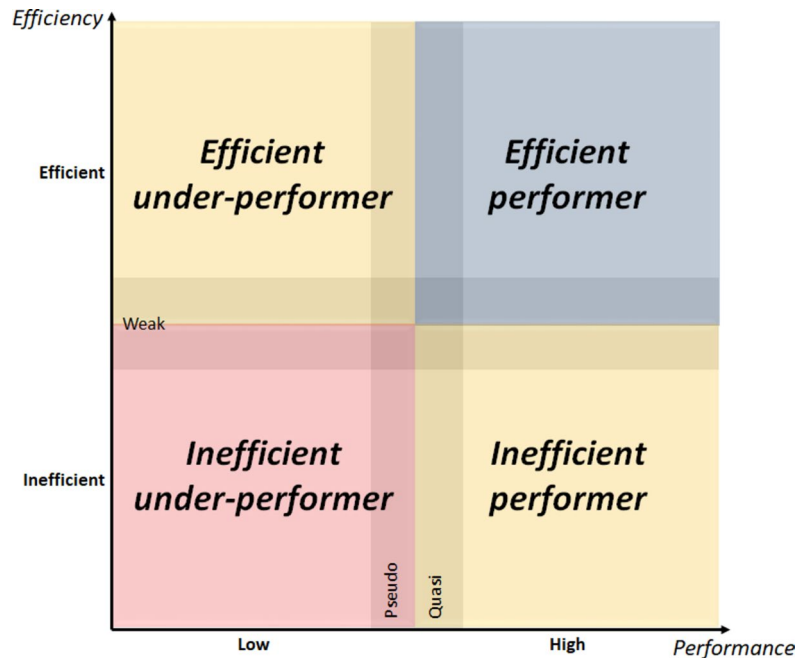


Figure 4. Qualitative bi-criteria visual matrix.

$$\begin{aligned}
 DMU_i \text{ is DEA efficient at time } t &\Leftrightarrow \theta^t(DMU_i) = 1 \\
 &\text{and } \lambda^t(DMU_i) = 0 \text{ for all inputs} \\
 DMU_i \text{ is DEA inefficient at time } t &\Leftrightarrow \theta^t(DMU_i) < 1 \\
 DMU_i \text{ is DEA weakly efficient at time } t &\Leftrightarrow \theta^t(DMU_i) = 1 \\
 &\text{and } \lambda^t(DMU_i) \neq 0 \text{ for some inputs.}
 \end{aligned} \tag{14}$$

Slacks $\lambda^t(DMU_i)$ and efficiency $\theta^t(DMU_i)$ are obtained by solving Eq. 4 at each time period t .

Results

Experimentation platform

The experimentation platform and collaborative environment used in this study were set up on the cloud, with resources provided by WestGrid and Compute Canada under the COVID-19 resource allocation fast track. A total of 80 virtual CPUs, 300 GBs of RAM, and 10 TBs of persistent storage were allocated to us on the Arbutus cloud. This platform includes an online operational research toolbox that can run various data analytic algorithms, accessible via the website: (<https://heacovid.uvic.ca/explorer/>). Additionally, the platform is designed to be flexible for different levels of analysis (e.g. micro and miso) and can be extended to other jurisdictions.

Data collection and descriptive statistics

We collected, consolidated, and combined data from multiple trusted secondary data sources including curated online databases. The data is structured according to three categories: preparedness for the pandemic, population vulnerability to COVID-19 and clinical outcomes. We retrieved data related to preparedness from the World Bank database¹⁰⁰, the number of intensive care unit (ICU) beds data from the Statista¹⁰¹ and healthcare expenditure per capita from the OECD dataset⁹. We gathered the number of populations suffering from asthma and chronic obstructive pulmonary disease from the Lancet Respiratory Medicine journal¹⁰², data on diabetes, smoking and cancer from the World Population Review¹⁰³, cardiac disease, vascular disease, obesity and hypertension from the European Heart Network website¹⁰⁴ and the data on chronic kidney disease was collected from¹⁰⁵. The urban density and statistics on the population above 65 years of age were collected from the World Bank database¹⁰⁶ and the Our World in Data website¹⁰⁷, respectively. Additional data on preparedness and population vulnerability were collected from^{108,109} and¹¹⁰. The COVID-19 related data was collected from Harvard University's Dataverse¹¹¹ and the WHO's Global research on coronavirus disease¹¹². Finally, we used the United Nations Educational, Scientific and Cultural Organization (UNESCO) database to retrieve countries' demographic and contextual data¹¹³.

In this study, we used more than 23,000 data points collected from various trusted data sources. Descriptive statistics for the 5 preparedness and 15 vulnerability dimensions of the 37 OECD countries are presented in Table 3. We collected time series data on confirmed COVID-19 death and recovery cases for each OECD country from January 14, 2020 to July 11, 2020. We observed that the first few days of each time series included high variability due to the low number of cases reported. Therefore, we truncated the time series for each country by

Preparedness dimension	Mean	SD	Vulnerability dimension	Mean	SD
X ₁	471	257	V ₁	268	107
X ₂	336	95	V ₂	398	121
X ₃	888	425	V ₃	6,463	2,217
X ₄	10	7	V ₄	6,209	1,858
X ₅	4,224	2,045	V ₅	5,018	1,484
			V ₆	2,598	803
			V ₇	27,115	6,324
			V ₈	18,949	5,055
			V ₉	17,505	4,241
			V ₁₀	24,947	7,412
			V ₁₁	24,866	5,461
			V ₁₂	23,805	6,255
			V ₁₃	306	57
			V ₁₄	10	9
			V ₁₅	720	392

Table 3. Descriptive statistics for the preparedness and population vulnerability to COVID-19 dimensions in the OECD countries (per 100,000 people).

removing the first few days corresponding to less than 100 cumulative confirmed cases. The first day of each time series is determined when the number of confirmed cases is at least 100. Consequently, day 1 of the pandemic is consistent for all OECD countries. Additional descriptive statistics are provided in tables: table S1, table S2, table S3, table S4 and table S5; in the supplementary document.

Clustering countries in homogeneous groups with similar vulnerability to COVID-19

We use a clustering method called Fuzzy J-Means (FJM) to classify national healthcare systems (NHCSS) into homogeneous groups based on their population's vulnerability to COVID-19 (see Table 2). The FJM method, similar to other clustering techniques requires the number of clusters in the data to be pre-specified (see [fuzzy clustering: controlling for the population vulnerability](#) in section. Fuzzy Clustering). As such, the performance of the clustering method may be affected by the chosen value of k . In this study, we employed an objective evaluation measure based on the objective function of the fuzzy clustering problem⁹⁶. After several runs, we found that $k = 5$ was the most relevant number of clusters. The FJM algorithm then produced the final clustering results for the OECD countries, as summarized in Table 4. Detailed computations of the membership degrees of each country to each cluster can be found in Table S6 in the supplementary document.

Figure 5 presents a visualization of the population vulnerability profiles of each country in each cluster based on the 15 dimensions. This figure serves as a graphical validation of the results obtained from the fuzzy clustering approach. It is evident that clusters 2 to 5 consist of countries with similar vulnerability profiles, while Chile and Japan (cluster 1) are outliers. As these two countries are not representative of the other clusters, they are not included in the further analysis to ensure the validity of any efficiency and performance analysis.

Additionally, it is also noteworthy that, besides the FJM algorithm, several other clustering algorithms were tested using our experimentation platform, including the k-Means and hierarchical algorithms. Ultimately, we found that the fuzzy clustering FJM algorithm provided more accurate results.

Clinical performances of OECD countries' NHCS in response to COVID-19

We used the results from the clustering to compare the clinical performance of NHCSS within the same group. Figures 6, 7, 8, and 9 show the comparison of countries' clinical performance within each cluster after 25 and 111 days of the pandemic being declared (i.e., after registering more than 100 cumulative confirmed cases). The pre-clustering figures are provided in Appendix B. The yellow dashed lines represent the average daily death

Cluster	OECD Countries
Cluster 1	Chile, Japan
Cluster 2	Australia, Colombia, Israel, Mexico, New Zealand, Portugal, Spain, United Kingdom (UK)
Cluster 3	Belgium, Denmark, Germany, Greece, Iceland, Ireland, Luxembourg, Netherlands, Norway, Sweden, United States of America (USA)
Cluster 4	Austria, Estonia, Finland, France, Italy, Latvia, Switzerland
Cluster 5	Canada, Czech Republic, Hungary, Lithuania, Poland, Slovakia, Slovenia, South Korea, Turkey

Table 4. Clustering results of countries based on their population vulnerabilities.

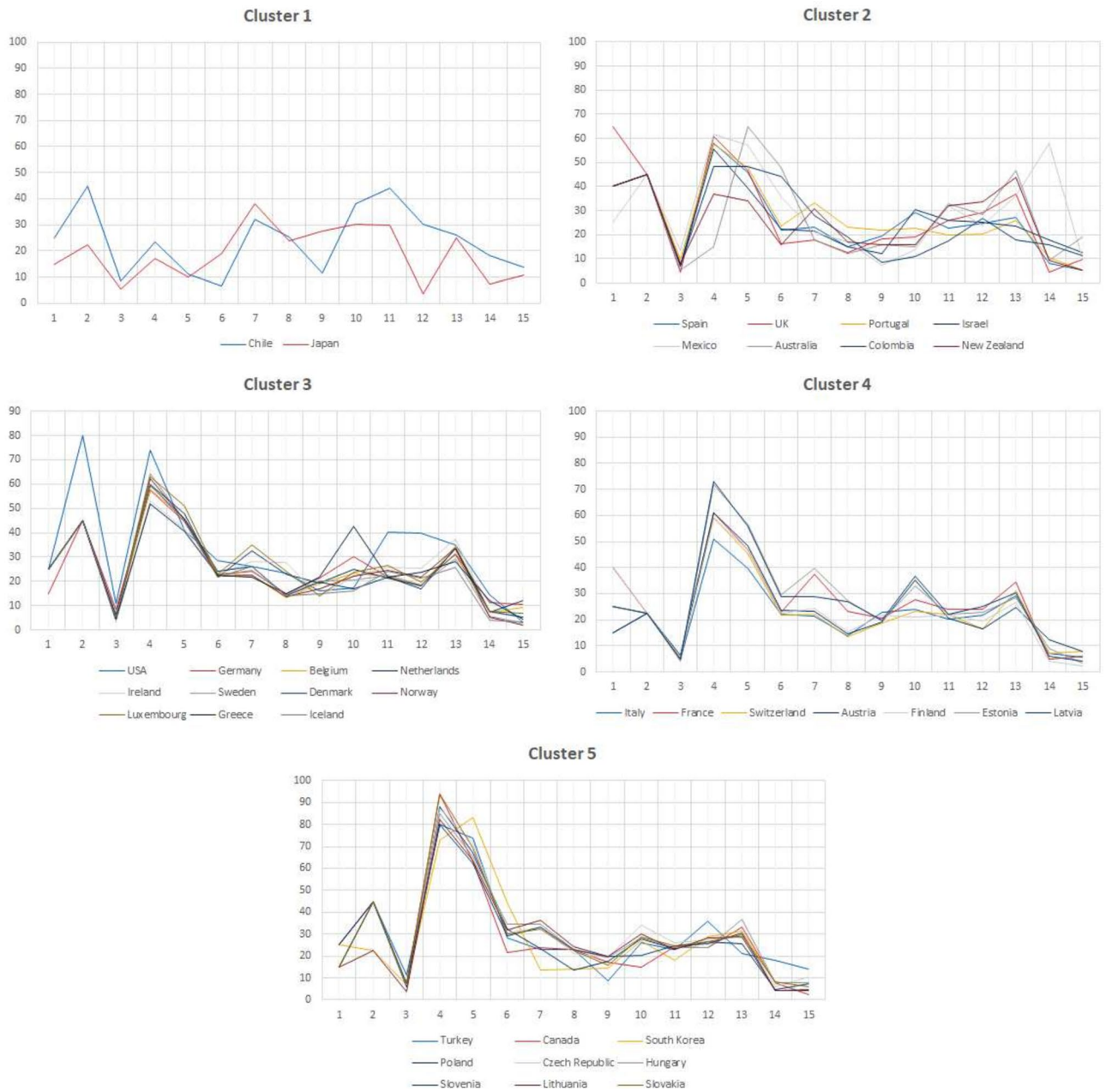


Figure 5. Mapping of population vulnerability profiles for each cluster.

and recovery rates for each cluster. For example, New Zealand is clearly a strong performer among its peers in cluster 2 after 111 days, while Spain is an under-performer. Germany is a leading performer in its group in cluster 3, while the USA is a quasi-performer. Both France and Italy under-perform their peers in cluster 4. Canada is an under-performer among its peers in cluster 5, while South Korea is a clear leader in the same group (Fig. 9).

Longitudinal efficiency of NHCS responses to the outbreak of COVID-19 pandemic

The results of our analysis using the DEA model for each cluster are shown in Fig. 10. This model compares the relative efficiency of countries in using their inputs to achieve outputs (for more details on DEA models, see longitudinal data envelopment analysis in section). In the first 100 days of the pandemic, the efficiency index for each country ranges from 0.30 to 1 (100%). With some exceptions, most countries generally improved their efficiency, though it remained low overall. Differences in efficiency averages between clusters are illustrated in Fig. 11. Clusters 2 and 5 have similar averages of 0.76 and 0.77, respectively, showing a loss of efficiency in the first 20 days of the pandemic, followed by gradual improvement. Clusters 3 and 4 have lower averages of 0.68 and 0.70, respectively, showing a loss of efficiency for over 40 days before improving.

Among the countries, New Zealand and South Korea stand out as leaders, demonstrating exemplary clinical outcomes and efficiency. Their success serves as a potential model for other countries to learn from. Spain, despite reaching the efficiency frontier early, maintained a stable efficiency slightly above 0.5. In cluster 3, most

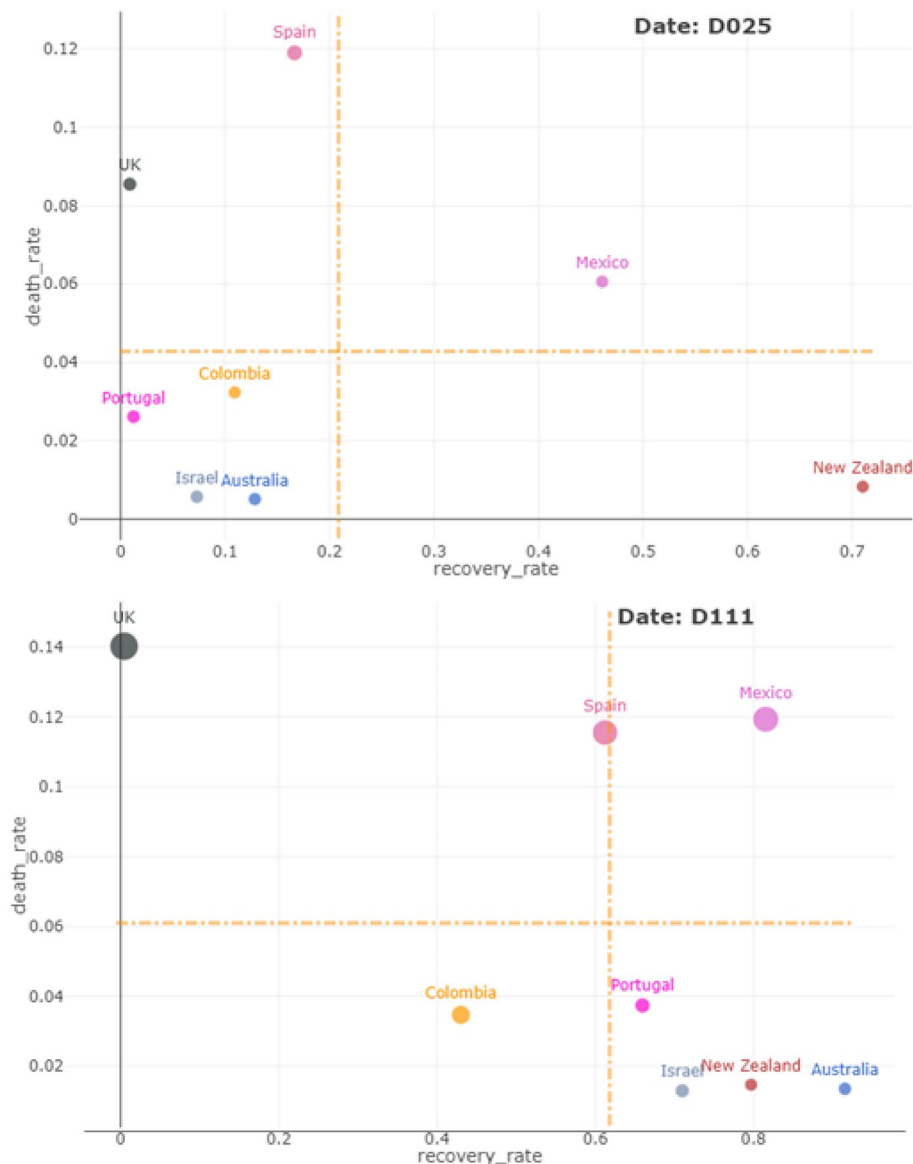


Figure 6. Clinical performance comparison among cluster 2 countries after 25 and 111 days of the pandemic.

countries struggled with unstable and low-efficiency levels, ranging between 0.34 and 1. The USA, initially on the efficiency frontier, experienced a drop to 0.63. Germany, despite its clinically solid performance, had an efficiency level of 1 for 14 days, then dropped to only 0.34 and then gradually improved, with an average of 0.66 due to under-utilized resources. France and Italy in cluster 4 had similar behaviours, achieving average efficiencies of 0.55 and 0.53, respectively. South Korea in cluster 5, however, reached the optimal efficiency frontier after 26 days, indicating effective use of NHCS resources. Canada, a consistent performer only at the beginning of the pandemic, achieved efficiency periodically and maintained stable efficiency levels of around 0.79 throughout the pandemic, with an average efficiency of 0.82.

The Malmquist index captures the changes in efficiency over time, as shown in Fig. 12. A value higher than 1 for the Malmquist index at time t indicates increased productivity. In contrast, a value less than 1 implies a decline in productivity compared to the previous period ($t - 1$). Spain and Italy exhibit similar patterns, with significant changes in productivity, likely due to being among the first countries to experience a high number of cases and not being fully prepared with the necessary resources to handle the pandemic. The Malmquist index exceeded 14 for Spain and reached 2.75 for Italy during the first fourteen days. This variation is primarily driven by an increase in the Malmquist institutional efficiency change (MTEC) index during this period, which measures the catch-up effect between two consecutive periods, ($t - 1$) and t . Conversely, the drop in productivity on day seven and the subsequent increase on day eight are attributed to fluctuations in the Malmquist technical change (MTC) index, which measures the frontier-shift effect or the change in the reference frontier between two consecutive periods ($t - 1$) and t , as shown in Fig. 13.

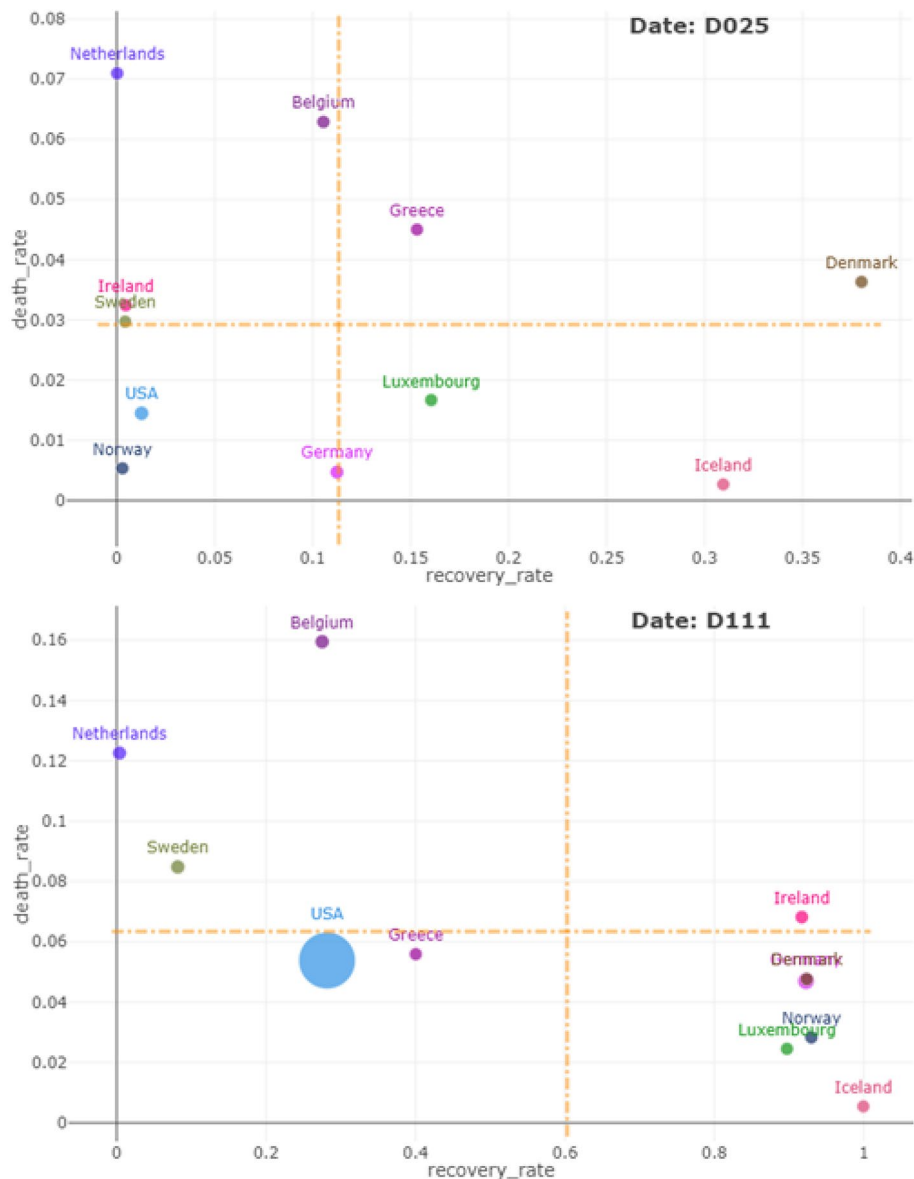


Figure 7. Clinical performance comparison among cluster 3 countries after 25 and 111 days of the pandemic.

Two weeks after the detection of the 100th case, productivity levels stabilized. Notably, France's productivity variation differs from Italy's and Spain's, remaining flat at around 1, indicating no change in the system's productivity. The productivity levels of New Zealand and South Korea also show noteworthy behaviour. South Korea had the slightest fluctuation in productivity, which may be due to its ability to manage and control the pandemic effectively. The USA's productivity averaged around 0.93, mainly due to a drop in the MTC (see Fig. 13). Canada's Malmquist index stabilized at 0.98; however, unlike the USA, which had maximum productivity of 1.05, Canada's productivity level reached as high as 2.24 before quickly dropping.

In some cases, productivity variation is due to data reporting inconsistencies rather than objective factors (e.g., most of the productivity gain in Mexico is caused by fluctuations in the number of reported cases). It is important to note that our analysis is based on publicly available data, which may have limitations in terms of accuracy and consistency. The Malmquist index is a complex metric that may only partially capture some aspects of productivity changes. These factors should be considered when interpreting our findings.

Bi-criteria comparison of NHCS responses to the outbreak of the COVID-19 pandemic over time

The results of the qualitative multi-criteria analysis for each country are shown in Fig. 14 (see [bi-criteria visual analysis of efficiency and performance of NHCS](#) in section and Algorithm 1). Figure 14 presents the efficiency-performance paths for selected countries from each cluster: New Zealand and Spain from Cluster 2, Germany and the USA from Cluster 3, Italy and France from Cluster 4, and Canada and South Korea from Cluster 5.

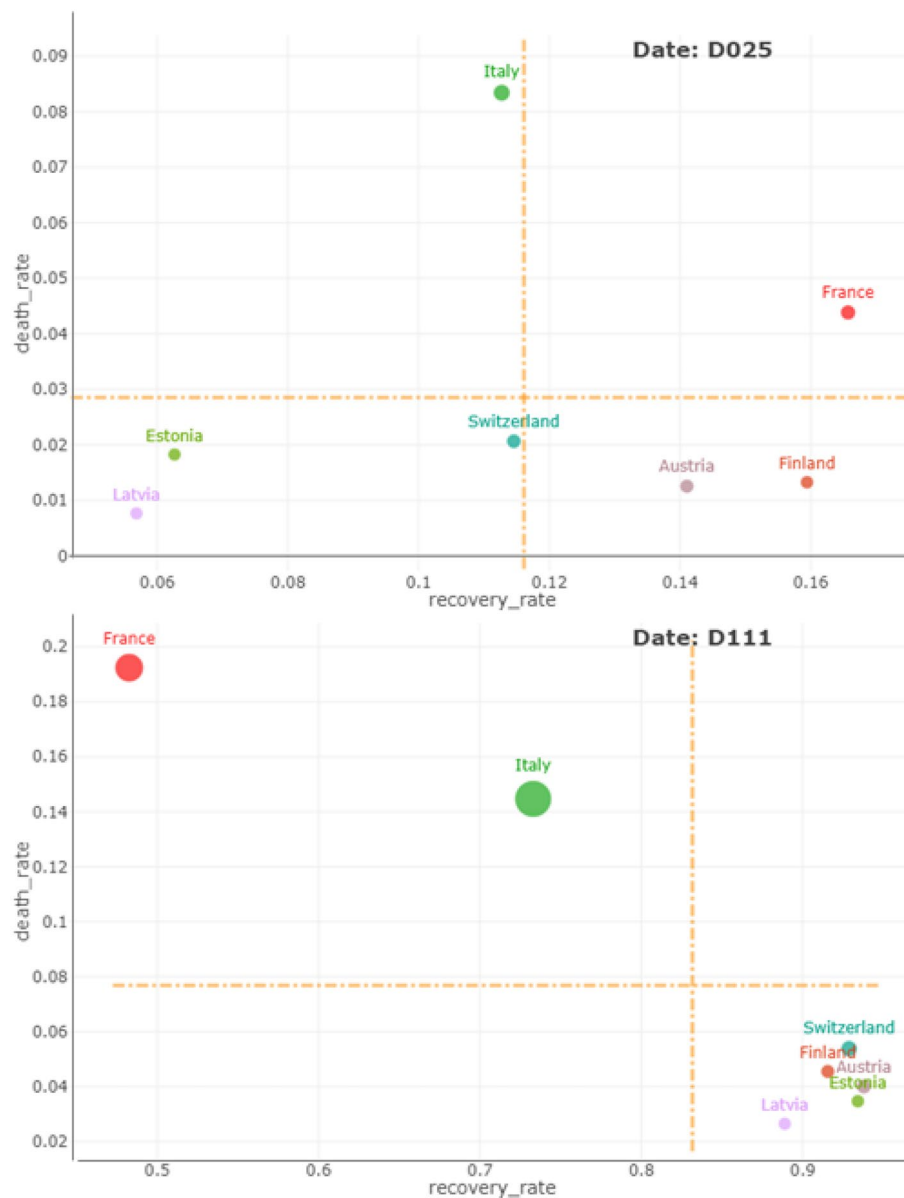


Figure 8. Clinical performance comparison among cluster 4 countries after 25 and 111 days of the pandemic.

Analyzing Canada's performance, we observe significant fluctuations. On day 1 of the pandemic, Canada effectively led the pack but quickly lost its edge. By day 13, it was considered a pseudo-performer among its peers in cluster 5 while simultaneously being on the efficiency frontier. However, by day 14, Canada had become an inefficient performer. On day 19, it regained pseudo-efficient performer status in cluster 5 but lost this advantage by day 25, becoming an under-performer and inefficient. This decline could be attributed to an increase in fatalities in long-term care homes. On the other hand, Germany began as an efficient under-performer and, by day 16 of the pandemic, had moved to the category of inefficient performers within its cluster. New Zealand consistently outperformed its peers, suggesting it should be studied for best practices. Additionally, Figure S1 in the supplementary document illustrates the dynamic nature of NHCS responses to the pandemic over time in each country. For more detailed information, readers are referred to the web platform (<https://heacovid.uvic.ca/explorer/>) to explore the qualitative multi-criteria map for each country and cluster.

Discussion and managerial implications

In this study, we present a novel approach to comparing national healthcare systems' (NHCS) responses to the COVID-19 pandemic by clearly defining the boundaries of the unit of analysis, which in our case refers to the entire healthcare system of a country. Our proposed model highlights the essential capabilities for effective response: acute and long-term care and protecting front-line workers.

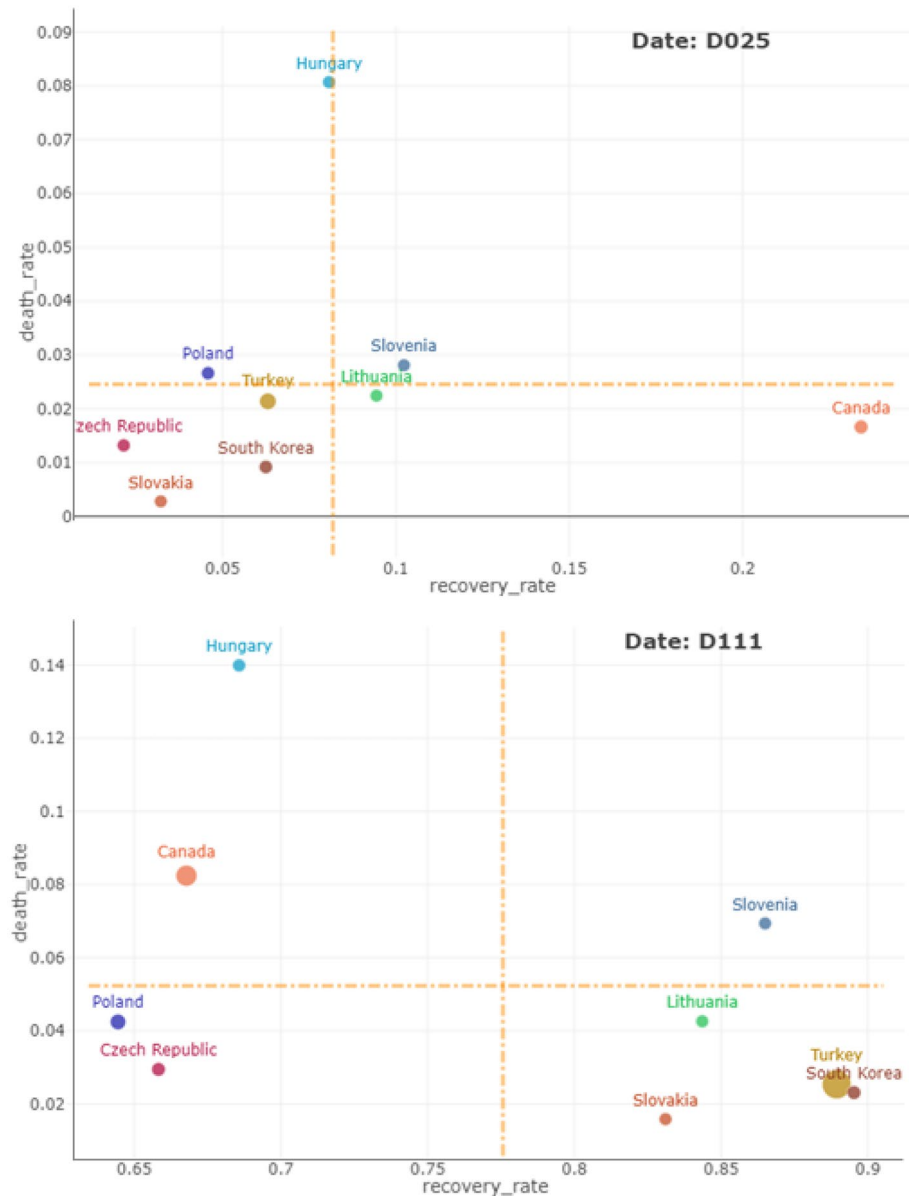


Figure 9. Clinical performance comparison among cluster 5 countries after 25 and 111 days of the pandemic.

We introduce a bi-criteria algorithm that integrates longitudinal DEA and fuzzy clustering, a method that allows for the classification of data points into clusters based on their similarity. This algorithm evaluates NHCS performance based on two critical criteria: efficiency, which measures the system's ability to deliver healthcare services with minimal resource wastage, and clinical outcomes, which assess the system's effectiveness in improving patient health. The study focuses on OECD countries due to the availability of reliable data, but the proposed methodology can be extended to other regions as data quality improves.

Our validation using OECD data revealed significant variability in NHCS responses, with several lessons emerging for improving future pandemic responses. The analysis confirms that OECD countries exhibited varied performance levels during the COVID-19 outbreak. Our 'mixed methodology', which combines quantitative and qualitative approaches, enables meaningful comparisons among healthcare systems serving similar populations by controlling for population vulnerability. Exemplary performers included New Zealand, Germany, Austria, and South Korea.

The study's results indicate that not all countries reached the relative efficiency frontier within the first 100 days of the pandemic, experiencing different performance dynamics. The combined efficiency and productivity variation analyses suggest that many healthcare systems must utilize resources optimally. Productivity gains were primarily driven by increased cases rather than efficient resource use, pointing to initial under-utilization. The premature closure of non-urgent healthcare services exacerbated inefficiencies and created backlogs, highlighting the need for more adaptive strategies.

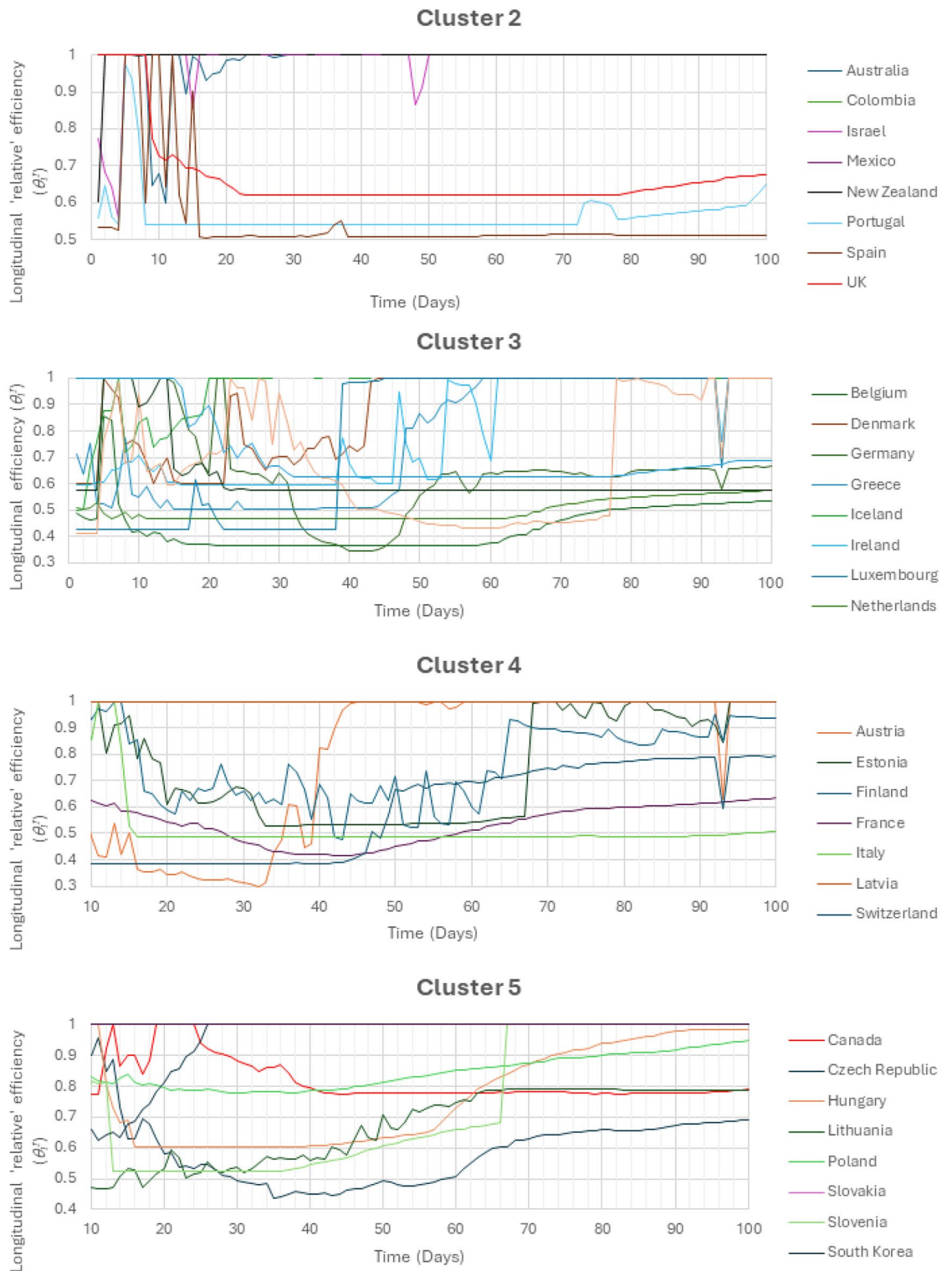


Figure 10. DEA longitudinal results for clusters 2 to 5.

Our findings also underscore disparities in protecting frontline workers. Effective measures in countries like South Korea, Japan, and Taiwan contrast sharply with struggles in the United States and the United Kingdom. This protection disparity underscores the critical importance of safeguarding healthcare workers during crises.

Regarding acute and long-term care, Germany, Austria, and New Zealand demonstrated effective responses, expanding intensive care units, increasing hospital bed capacity, and ensuring adequate staffing. These measures facilitated better patient influx management and underscored the importance of scalable healthcare infrastructure.

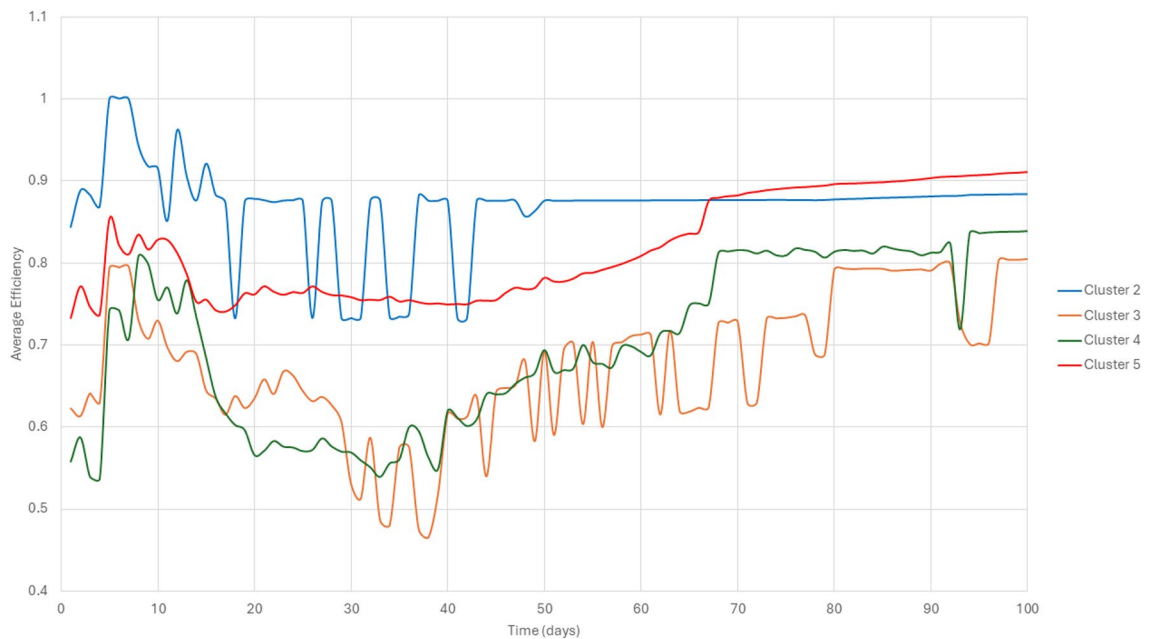


Figure 11. Longitudinal average efficiency index for each cluster.

The study's findings emphasize the necessity of rethinking pandemic response strategies and redesigning healthcare service delivery in alignment with public health policies. However, these findings also present a unique opportunity for improvement. By implementing the recommendations outlined in this study, healthcare systems can enhance their preparedness and response to future pandemics. For instance, adaptive strategies could include the establishment of flexible healthcare facilities that can be quickly converted into isolation wards during a pandemic, or the development of dynamic staffing plans that can be adjusted based on the severity of the outbreak. These examples inspire a sense of optimism and motivation in the audience.

We propose the following critical managerial implications:

- Healthcare systems should adopt dynamic and collaborative forecasting and planning approaches with rolling planning horizons to anticipate and respond to changing demands effectively.
- Efficient resource utilization is crucial. Systems should avoid abrupt service cancellations and adopt 'phased approaches' to manage capacity without compromising other healthcare services. For example, hospitals could prioritize urgent surgeries and procedures during the pandemic's peak while gradually resuming non-urgent services as the situation improves.
- Protecting front-line workers should be a priority, with adequate personal protective equipment (PPE), mental health support, and clear protocols to ensure their safety and well-being. Several countries registered high numbers of fatalities in their frontline work force.
- Investment in scalable healthcare infrastructure, which refers to the ability of a healthcare system to rapidly expand its capacity in response to increased demand, including ICU capacity, hospital beds, and staffing, is essential to effectively managing patient surges during pandemics.
- Aligning healthcare delivery with public health policies ensures cohesive and coordinated responses. This includes integrating public health strategies with healthcare operations to manage outbreaks comprehensively.

Conclusion

Our study introduces a novel approach, a bi-criteria framework, to understand the performance dynamics of national healthcare systems (NHCSs) during the COVID-19 pandemic. This framework, which integrates DEA and fuzzy clustering, provides a comprehensive and dynamic assessment of NHCS performance. By applying this methodology, we uncover significant variability in NHCS responses among OECD countries and identify key factors influencing efficiency and performance.

Our findings carry significant practical implications, revealing that not all countries reached the relative efficiency frontier within the first 100 days of the pandemic. Many experienced fluctuating performance dynamics. Exemplary performers such as New Zealand, Germany, Austria, and South Korea demonstrated effective resource utilization and strong clinical outcomes. Conversely, countries like Spain and Italy exhibited significant initial productivity gains followed by declines, underscoring the challenges of sustaining effective responses. The disparities in protecting frontline workers and managing acute and long-term care highlight critical areas for improvement.

Implementing adaptive strategies, efficient resource utilization, data-driven decision-making, international collaboration, targeted support for vulnerable populations, and continuous training will enhance the resilience and effectiveness of healthcare systems worldwide. These recommendations are not just words on paper, they inspire hope and motivation in the audience, guiding future research and policy decisions.

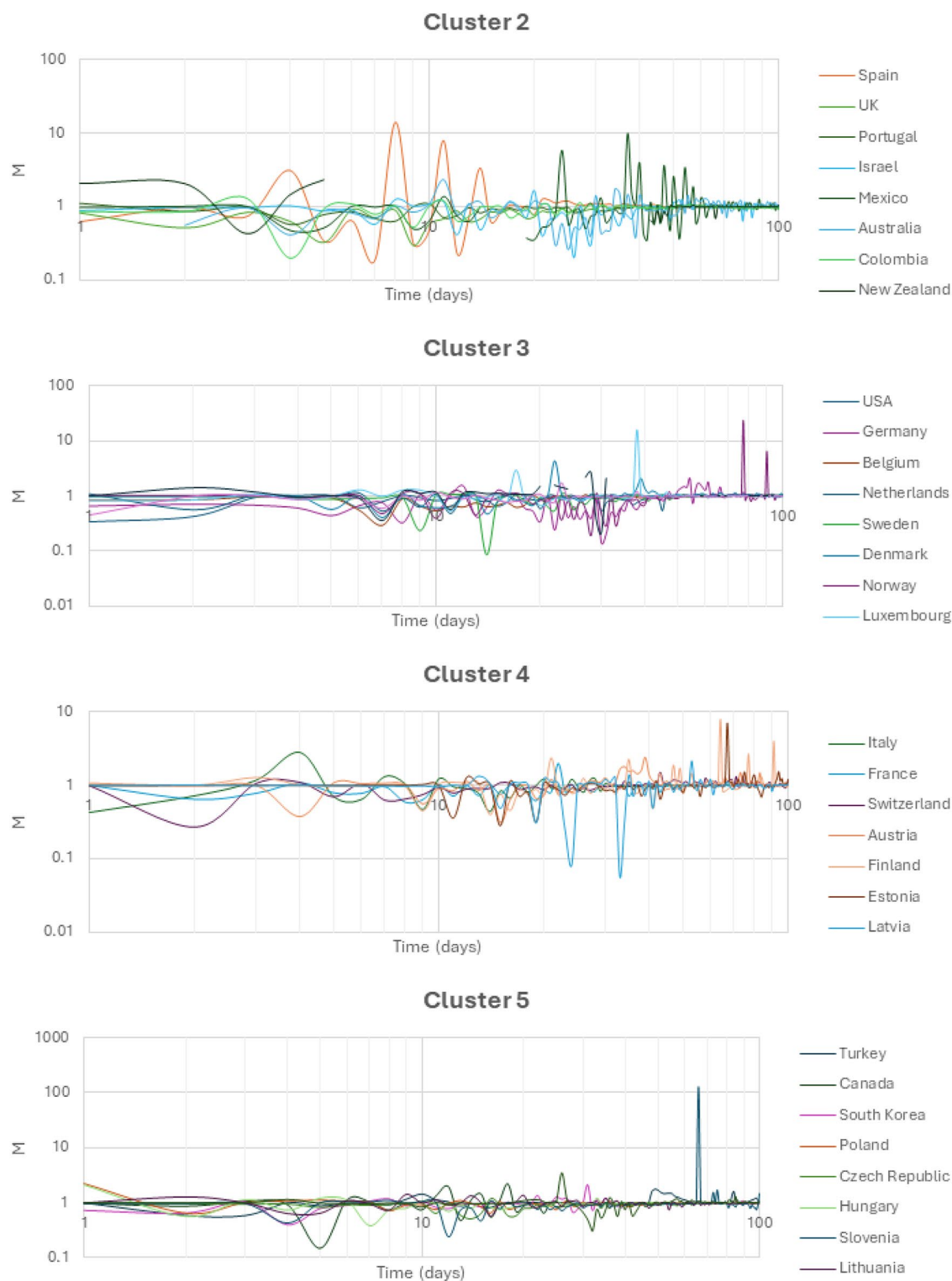


Figure 12. Malmquist index for each cluster.

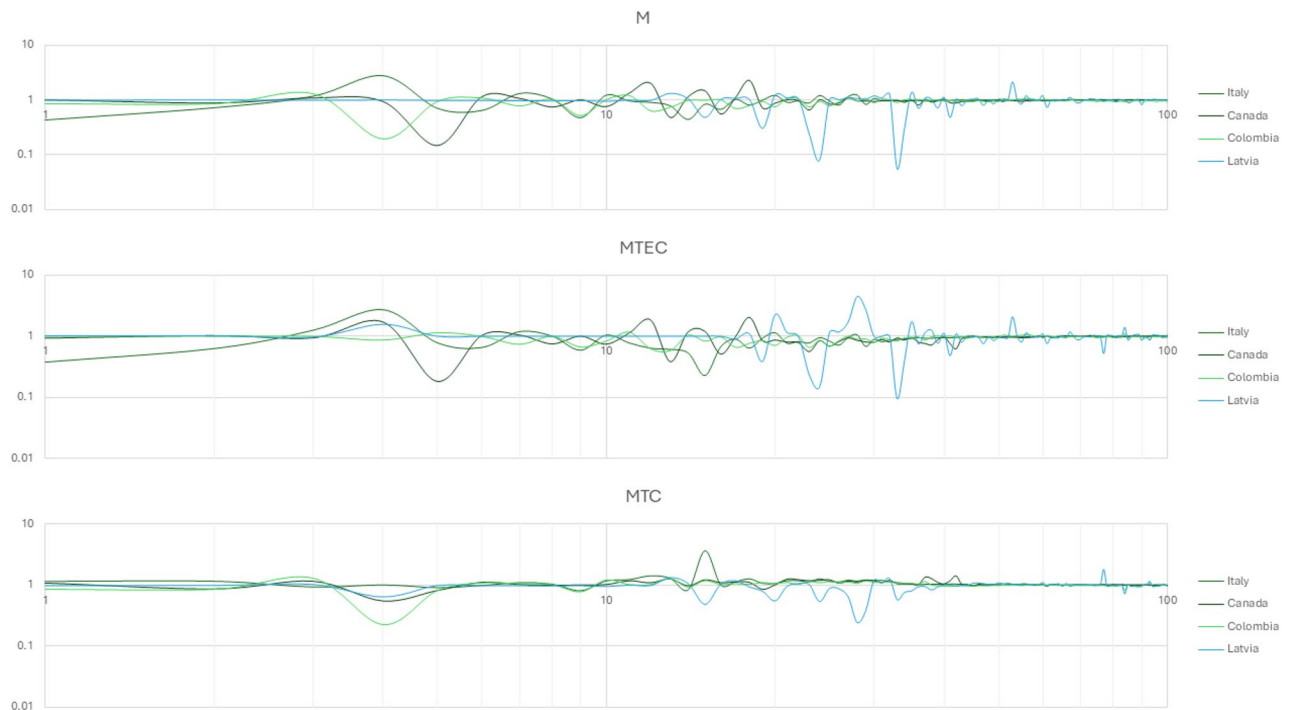


Figure 13. Malmquist index decomposition for selected countries.

While our bi-criteria framework effectively addresses the research questions and controls for moderating factors such as population vulnerabilities, there are inherent limitations. The robustness of the results heavily depends on the quality and consistency of data from various sources, which can vary significantly. Additionally, the reliance on ratios rather than absolute values in our DEA model was driven by data availability constraints. Although ratios facilitate comparisons across different health systems, they can obscure scale and absolute performance differences, potentially affecting the precision of efficiency assessments.

Future research should prioritize obtaining comprehensive absolute value data to enhance the robustness and accuracy of DEA models, enabling more direct and meaningful comparisons. Refining analytical methods, incorporating a more comprehensive range of criteria, and exploring advanced machine-learning techniques will improve predictive accuracy and robustness. Expanding the scope to include non-OECD countries will enable broader applicability and more comprehensive comparisons.

Despite these limitations, our research confirms that some NHCSs outperform others, highlighting the need for continuous evaluation and improvement of pandemic response strategies. This process demands ongoing commitment and collaboration among policymakers, healthcare managers, and researchers. By addressing these limitations and building on the findings of this study, future research can contribute significantly to the development of more resilient and effective healthcare systems.

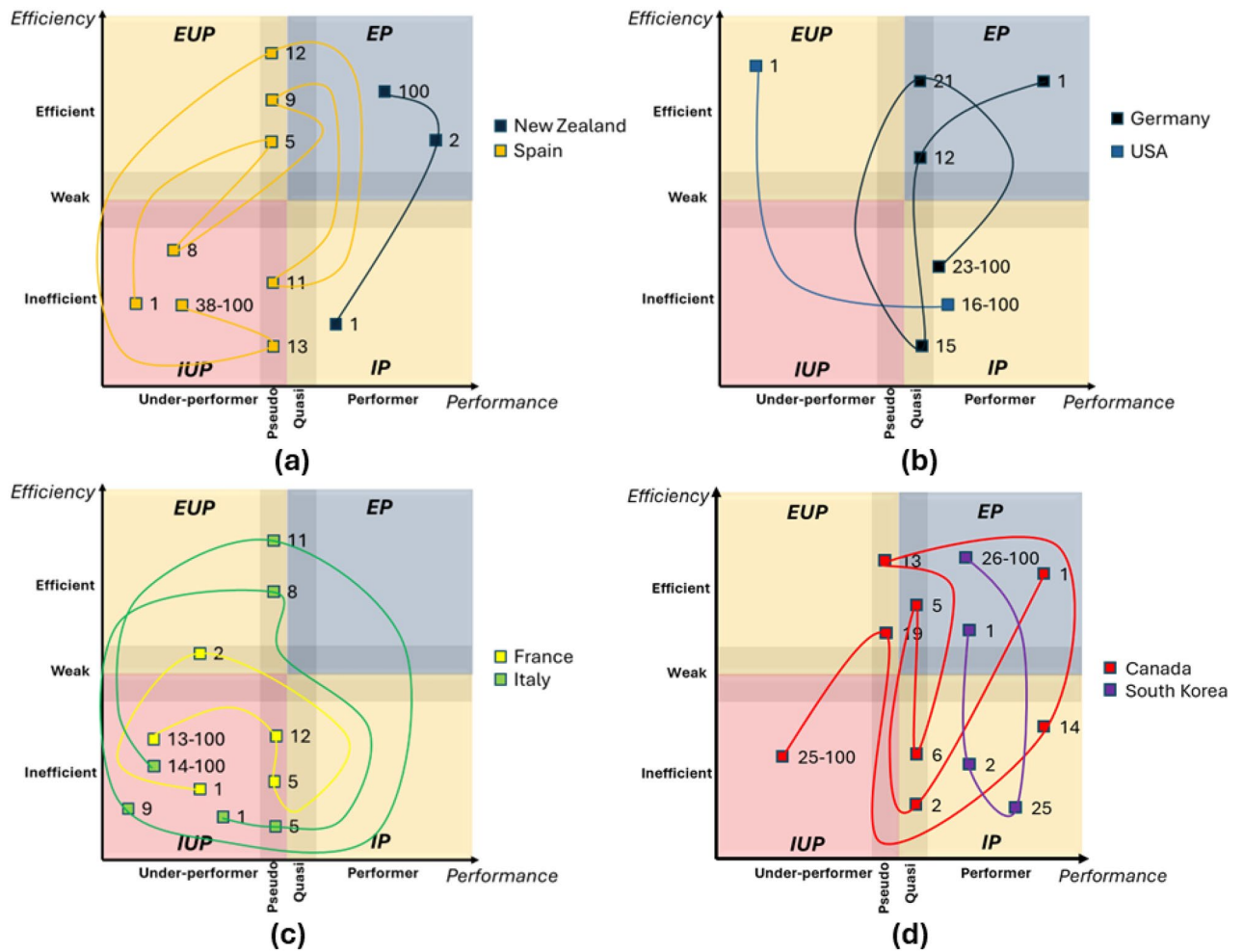


Figure 14. Qualitative analysis for selected countries: (a) New Zealand's NHCS maintained good performance over time but Spain's was an under performer in cluster 2, (b) Germany's NHCS outperforms USA's in cluster 3, (c) Italy's NHCS and France's seems to follow similar patterns in cluster 4, and (d) South Korea's NHCS seems to outperform Canada's in cluster 5.

Data availability

The dataset used in the current study are available from the corresponding author on reasonable request. All data is available through the project's web platform. All code and algorithms are available through the project's web platform and available upon request. This platform is available at <https://heacovid.uvic.ca/explorer/>. The fuzzy clustering FJM is available at <https://github.com/nbelacel/FuzzyJMeansVNS>.

Received: 20 February 2023; Accepted: 1 August 2024

Published online: 27 September 2024

References

- Ioannidis, J. P. A. Coronavirus disease 2019: The harms of exaggerated information and non-evidence-based measures. *Eur. J. Clin. Invest.* **50**, e13222–e13222. <https://doi.org/10.1111/eci.13222> (2020).
- Livingston, E., Bucher, K. & Rekito, A. Coronavirus disease 2019 and influenza 2019–2020. *JAMA* **323**, 1122–1122 (2020).
- Wang, L. *et al.* Real-time estimation and prediction of mortality caused by covid-19 with patient information based algorithm. *Sci. Total Environ.* **138394** (2020).
- Armocida, B., Formenti, B., Ussai, S., Palestra, F. & Missoni, E. The italian health system and the covid-19 challenge. *The Lancet Public Health* **5**, e253 (2020).
- Ellis-Petersen, H. The system has collapsed: India's descent into covid hell. *The Guardian* (2021).
- Lal, A., Arondu, N. A., Heymann, D. L., Gitahi, G. & Yates, R. Fragmented health systems in covid-19: rectifying the misalignment between global health security and universal health coverage. *The Lancet* (2020).
- Allen, W. E. *et al.* Population-scale longitudinal mapping of covid-19 symptoms, behavior, and testing identifies contributors to continued disease spread in the united states. *medRxiv* (2020).
- WHO. 2019 novel coronavirus (2019-ncov): strategic preparedness and response plan. *WHO* (2020).
- Muller, J. *et al.* Major interregional differences in france of covid-19 hospitalization and mortality from january to June 2020. *Revue D'épidémiologie et de Santé Publique* **70**, 265–276. <https://doi.org/10.1016/j.respe.2022.08.008> (2022).
- Papanicolas, I. & Smith, P. *Health system performance comparison: an agenda for policy, information and research: an agenda for policy, information and research* (McGraw-Hill Education UK., 2013).

11. Roberts, M. J., Hsiao, W., Berman, P. & Reich, M. R. Getting health reform right: a guide to improving performance and equity. *New York* (2008).
12. WHO. Everybody's business—strengthening health systems to improve health outcomes: Who's framework for action. *who* (2007).
13. Flood, A. B. & Fennell, M. L. Through the lenses of organizational sociology: the role of organizational theory and research in conceptualizing and examining our health care system. *Journal of health and social behavior* 154–169 (1995).
14. Nembhard, I. M. *et al.* Moving organizational theory in health care forward: A discussion with suggestions for critical advancements. *Health Care Manag. Rev.* **45**, E1–E12 (2020).
15. Murray, C. J. & Frenk, J. A framework for assessing the performance of health systems. *Bull. world Health Organ.* **78**, 717–731 (2000).
16. Anderson, L. M. *et al.* Culturally competent healthcare systems: A systematic review. *Am. J. Prevent. Med.* **24**, 68–79 (2003).
17. Zare, H., Tavana, M., Mardani, A., Masoudian, S. & Saraji, M. K. A hybrid data envelopment analysis and game theory model for performance measurement in healthcare. *Health Care Manag. Sci.* **22**, 475–488 (2019).
18. Grigoroudis, E., Orfanoudaki, E. & Zopounidis, C. Strategic performance measurement in a healthcare organisation: A multiple criteria approach based on balanced scorecard. *Omega* **40**, 104–119 (2012).
19. Lessons to be learnt from efficient countries. Muhammed Ordu, E. D., Hediye Kirli Akin. Healthcare systems and covid-19. *Int J Health Plann Mgmt* **36**, 1476–1485 (2021).
20. Pourmahmoud, J. & Bagheri, N. Uncertain malmquist productivity index: An application to evaluate healthcare systems during covid-19 pandemic. *Socio-Economic Planning Sciences* **87**, 101522 (2023).
21. Henriques, C. & Gouveia, M. Assessing the impact of covid-19 on the efficiency of portuguese state-owned enterprise hospitals. *Socio-economic planning sciences* **84**, 101387 (2022).
22. Lupu, D. & Tiganasu, R. Covid-19 and the efficiency of health systems in europe. *Health Economics Review* **12**, 14 (2022).
23. Mourad, N., Habib, A. & Tharwat, A. Appraising healthcare systems' efficiency in facing covid-19 through data envelopment analysis. *Decision Science Letters* **10**, 301–310 (2021).
24. Md Hamzah, N., Yu, M.-M. & See, K. F. Assessing the efficiency of malaysia health system in covid-19 prevention and treatment response. *Health Care Manag. Sci.* **24**, 273–285 (2021).
25. Cooper, W. W., Seiford, L. M. & Zhu, J. *Handbook on data envelopment analysis*, vol. 164 (Springer, 2011).
26. Lancet. Covid-19: protecting health-care workers. *Lancet* **395**, 922 (2020).
27. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: a retrospective cohort study. *The Lancet* (2020).
28. Banerjee, A. *et al.* Estimating excess 1-year mortality associated with the covid-19 pandemic according to underlying conditions and age: A population-based cohort study. *The Lancet* **395**, 1715–1725. [https://doi.org/10.1016/S0140-6736\(20\)30854-0](https://doi.org/10.1016/S0140-6736(20)30854-0) (2020).
29. Onder, G., Rezza, G. & Brusaferro, S. Case-fatality rate and characteristics of patients dying in relation to covid-19 in italy. *JAMA* **323**, 1775–1776 (2020).
30. Puelles, V. G. *et al.* Multiorgan and renal tropism of sars-cov-2. *New England Journal of Medicine* (2020).
31. Shi, S. *et al.* Characteristics and clinical significance of myocardial injury in patients with severe coronavirus disease 2019. *Eur. Heart J.* **41**, 2070–2079 (2020).
32. Zaki, N., Alashwal, H. & Ibrahim, S. Association of hypertension, diabetes, stroke, cancer, kidney disease, and high-cholesterol with covid-19 disease severity and fatality: A systematic review. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* (2020).
33. Yang, Q. *et al.* Effect of hypertension on outcomes of adult inpatients with covid-19 in wuhan, china: A propensity score-matching analysis. *Respir. Res.* **21**(171), 1–9 (2020).
34. Tan, Y.-K. *et al.* Covid-19 and ischemic stroke: a systematic review and meta-summary of the literature. *J. Thrombosis Thromb.* 1–9 (2020).
35. Mazucanti, C. H. & Egan, J. M. Sars-cov-2 disease severity and diabetes: Why the connection and what is to be done?. *Immun. Age.* **17**, 21 (2020).
36. Devaraj, S., Ow, T. T. & Kohli, R. Examining the impact of information technology and patient flow on healthcare performance: A theory of swift and even flow (tsef) perspective. *J. Oper. Manag.* **31**, 181–192. <https://doi.org/10.1016/j.jom.2013.03.001> (2013).
37. Ozcan, Y. A. *Evaluation of Performance in Health Care*, 3–14 (Springer, 2014).
38. Papanicolas, I. & Cylus, J. *Comparison of Healthcare Systems Performance*, chap. The Palgrave International Handbook of Healthcare Policy and Governance, 116–132 (Palgrave Macmillan, 2015).
39. Davis, K. *et al.* Mirror, mirror on the wall: An international update on the comparative performance of american health care. *New York Commonwealth Fund* **59** (2007).
40. Jamison, D. T. & Sandbu, M. E. Who ranking of health system performance. *Science* **293**, 1595–1596. <https://doi.org/10.1126/science.1059029> (2001).
41. Ozcan, Y. A. *Performance Measurement Using Data Envelopment Analysis (DEA)*, 15–47 (Springer, 2014).
42. Sow, O., Oukil, A., NDIAYE, B. & MARCOS, A. Efficiency analysis of public transportation subunits using dea and bootstrap approaches—dakar dem dikk case study. *J. Math. Res.* (2016).
43. Oukil, A. *et al.* A dea cross-efficiency inclusive methodology for assessing water quality: A composite water quality index. *J. Hydrol.* **612**, <https://doi.org/10.1016/j.jhydrol.2022.128123> (2022). Publisher Copyright: © 2022 Elsevier B.V.
44. Banker, R. D., Charnes, A. & Cooper, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* **30**, 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078> (1984).
45. Hassan, M. M. & Oukil, A. Design of efficient systems of commercial material handling equipment for supply chain and logistics facilities using DEA. *Int. J. Log. Syst. Manag.* **39**, 241–272 (2021).
46. Ozcan, Y. A. *Health Care Benchmarking and Performance Evaluation* (Springer, 2014).
47. Özgen Narci, H., Ozcan, Y. A., Şahin, İ., Tarcan, M., & Narci, M. examination of competition and efficiency for hospital industry in turkey. *Health Care Manag. Sci.* **18**, 407–418 (2015).
48. Hu, H.-H., Qi, Q. & Yang, C.-H. Analysis of hospital technical efficiency in China: Effect of health insurance reform. *China Econ. Rev.* **23**, 865–877 (2012).
49. Gautam, S., Hicks, L., Johnson, T. & Mishra, B. Measuring the performance of critical access hospitals in missouri using data envelopment analysis. *J. Rural Health* **29**, 150–158 (2013).
50. Aydin Nezir, G. Y. Assessing countries' performances against covid-19 via wsidea and machine learning algorithms. *Appl. Soft comput.* **97** (2020).
51. M.A Kamel, M. M. Measuring operational efficiency of isolation hospitals during covid-19 pandemic using data envelopment analysis: a case of egypt. *Benchmarking Int. J.* **ahead-of-print** (2021).
52. D.T. Jamison, K. W., L.J. Lau & Xiong, Y. Country performance against covid-19: rankings for 35 countries. *BMJ Global Health* **5** (2020).
53. M.I. Doğan, Özsoy, V. S. & Örkücü, H. H. Performance management of oecd countries on covid-19 pandemic: a criticism using data envelopment analysis models. *J. Facilities Manag.* **19**, 479–499 (2021).
54. Ranney, M. L., Griffeth, V. & Jha, A. K. Critical supply shortages—the need for ventilators and personal protective equipment during the covid-19 pandemic. *New Engl. J. Med.* **382**, e41 (2020).

55. Nicola, M. *et al.* The socio-economic implications of the coronavirus pandemic (covid-19): A review. *Int. J. Surg.* **78**, 185–193 (2020).
56. Greco, S., Figueira, J. & Ehrgott, M. *Multiple criteria decision analysis* Vol. 37 (Springer, 2016).
57. Wątróbski, J., Jankowski, J., Ziemia, P., Karczmarczyk, A. & Ziolo, M. Generalised framework for multi-criteria method selection. *Omega* **86**, 107–124 (2019).
58. Oukil, A. & Govindaluri, S. M. A hybrid multi-attribute decision-making procedure for ranking project proposals: A historical data perspective. *Manag. Dec. Econ.* **41**, 461–472. <https://doi.org/10.1002/mde.3113> (2020).
59. Oukil, A. & Amin, G. An induced owa aggregation operator with dual preference setting for dea cross-efficiency ranking. *Soft Comput.* **27**, 18419–18440. <https://doi.org/10.1007/s00500-023-09235-0> (2023).
60. Belton, V. & Stewart, T. *Multiple Criteria Decision Analysis: An Integrated Approach* (Springer, 2002).
61. Guitouni, A. & Martel, J.-M. Tentative guidelines to help choosing an appropriate mcda method. *Eur. J. Oper. Res.* **109**, 501–521 (1998).
62. Burton-Jones, A. *et al.* Changing the conversation on evaluating digital transformation in healthcare: Insights from an institutional analysis. *Inf. Organ.* **30**, 100255. <https://doi.org/10.1016/j.infoandorg.2019.100255> (2020).
63. Lawton, M. Environment and aging. center for the study of aging. *Albany, New York* 137 (1986).
64. Oukil, A. Selecting material handling equipment through a market weight scheme based dea cross-efficiency approach. *Int. J. Manag. Sci. Eng. Manag.* **19**, 1–14 (2024).
65. Gostin, L. O., Viswanathan, K., Altevogt, B. M. & Hanfling, D. *Crisis Standards of Care: A Systems Framework for Catastrophic Disaster Response: Volume 1: Introduction and CSC Framework*, vol. 3 (National Academies Press, 2012).
66. Hick, J. L., Barbera, J. A. & Kelen, G. D. Refining surge capacity: Conventional, contingency, and crisis capacity. *Disaster Med Public Health Prep* **3**, S59–67. <https://doi.org/10.1097/DMP.0b013e31819f1ae2> (2009).
67. Tang, Y.-W., Schmitz, J. E., Persing, D. H. & Stratton, C. W. Laboratory diagnosis of covid-19: current issues and challenges. *J. Clin. Microbiol.* **58** (2020).
68. Cavallo, J. J., Donoho, D. A. & Forman, H. P. Hospital capacity and operations in the coronavirus disease 2019 (covid-19) pandemic-planning for the nth patient. In *JAMA Health Forum* (American Medical Association, 2020).
69. WHO. Toolkit for assessing health-system capacity for crisis management. *Copenhagen: WHO* (2012).
70. Sonenthal, P. D. *et al.* Covid-19 preparedness in malawi: a national facility-based critical care assessment. *The Lancet Global Health* (2020).
71. Eissa, N. Pandemic preparedness and public health expenditure. *Economies* **8**, 60 (2020).
72. Banker, R. D., Charnes, A. & Cooper, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* **30**, 1078–1092 (1984).
73. Oukil, A., Kennedy, R., Al-Hajri, A. & Soltani, A. Unveiling the potential of hotel mergers: A hybrid dea approach for optimizing sector-wide performance in the hospitality industry. *International Journal of Hospitality Management* **116**, <https://doi.org/10.1016/j.ijhm.2023.103620> (2024). Publisher Copyright: © 2023 Elsevier Ltd.
74. Emrouznejad, A. & Yang, G.-L. A survey and analysis of the first 40 years of scholarly literature in dea: 1978–2016. *Socio-Econ. Plan. Sci.* **61**, 4–8 (2018).
75. Oukil, A. *et al.* A surrogate water quality index to assess groundwater using a unified dea-owa framework. *Environ. Sci. Pollut. Res. Int.* **28**, <https://doi.org/10.1007/s11356-021-13758-0> (2021).
76. Wang, S. & Sun, X. The global system-ranking efficiency model and calculating examples with consideration of the nonhomogeneity of decision-making units. *Exp. Syst.* **37**, e12272 (2020).
77. Thanassoulis, E. A data envelopment analysis approach to clustering operating units for resource allocation purposes. *Omega* **24**, 463–476 (1996).
78. Dai, X. & Kuosmanen, T. Best-practice benchmarking using clustering methods: Application to energy regulation. *Omega* **42**, 179–188 (2014).
79. Cinaroglu, S. Integrated k-means clustering with data envelopment analysis of public hospital efficiency. *Health Care Manag. Sci.* **23**, 325–338. <https://doi.org/10.1007/s10729-019-09491-> (2020).
80. Omrani, H., Shafaat, K. & Emrouznejad, A. An integrated fuzzy clustering cooperative game data envelopment analysis model with application in hospital efficiency. *Exp. Syst. Appl.* **114**, 615–628 (2018).
81. An analytical framework. M. Zarrin, J. S. & Brunner, J. Homogeneity and best practice analyses in hospital performance management. *Health Care Manag. Sci.* **25**, 406–425 (2022).
82. Moghaddas, Z., Oukil, A. & Vaez-Ghasemi, M. Global multi-period performance evaluation-new model and productivity index. *RAIRO-Oper. Res.* **56**, 1503–1521 (2022).
83. Hajihassaniasl, S. A comparative analysis of efficiency and productivity changes in turkish banking: A malmquist data envelopment analysis approach. *Int. J. Prod. Qual. Manag.* **6**, 383–397 (2021).
84. Färe, R. & Grosskopf, S. Directional distance functions and slacks-based measures of efficiency: Some clarifications. *Eur. J. Oper. Res.* **206**, 702–702 (2010).
85. Álvarez, I. *et al.* *A data envelopment analysis toolbox for matlab* (Universidad Autónoma de Madrid (Spain), Department of Economic Analysis, Tech. Rep., 2016).
86. Hollingsworth, B. Non-parametric and parametric applications measuring efficiency in health care. *Health Care Manag. Sci.* **6**, 203–18. <https://doi.org/10.1023/a:1026255523228> (2003).
87. Chhiba, K. D. *et al.* Prevalence and characterization of asthma in hospitalized and nonhospitalized patients with covid-19. *J. Allergy Clin. Immunol.* **146**, 307–314.e4. <https://doi.org/10.1016/j.jaci.2020.06.010> (2020).
88. Kang, S. *et al.* Recent progress in understanding 2019 novel coronavirus (sars-cov-2) associated with human respiratory disease: detection, mechanisms and treatment. *Int. J. Antimicrobial Agents* **55**, 105950–105950. <https://doi.org/10.1016/j.ijantimicag.2020.105950> (2020).
89. Zhao, Q. *et al.* The impact of copd and smoking history on the severity of covid-19: a systemic review and meta-analysis. *J. Med. Virol.* (2020).
90. Cappuccio, F. P. & Siani, A. Covid-19 and cardiovascular risk: Susceptibility to infection to sars-cov-2, severity and prognosis of covid-19 and blockade of the renin-angiotensin-aldosterone system. an evidence-based viewpoint. *Nutrition, metabolism, and cardiovascular diseases : NMCD* **30**, 1227–1235. <https://doi.org/10.1016/j.numecd.2020.05.013> (2020).
91. Iba, T., Levy, J. H., Levi, M., Connors, J. M. & Thachil, J. Coagulopathy of coronavirus disease 2019. *Crit. Care Med.* **48**, 1358–1364. <https://doi.org/10.1097/CCM.0000000000004458> (2020).
92. Dugail, I., Amri, E.-Z. & Vitale, N. High prevalence for obesity in severe covid-19: Possible links and perspectives towards patient stratification. *Biochimie* (2020).
93. Petrakis, D. *et al.* Obesity-a risk factor for increased covid-19 prevalence, severity and lethality. *Mol. Med. Rep.* **22**, 9–19 (2020).
94. WHO. Report of the who-china joint mission on coronavirus disease 2019 (covid-19). Available at <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf> (July 21,2020) (2020).
95. Rocklöv, J. & Sjödin, H. High population densities catalyse the spread of covid-19. *J. Travel Med.* **27**, taaa038 (2020).
96. Belacel, N., Hansen, P. & Mladenovic, N. Fuzzy j-means: A new heuristic for fuzzy clustering. *Pattern Recogn.* **35**, 2193–2200 (2002).

97. Belacel, N., Cuperlovic-Culf, M., Laflamme, M. & Ouellette, R. Fuzzy j-means and vns methods for clustering genes from microarray data. *Bioinformatics* **20**, 1690–1701 (2004).
98. Shakirova, A. *et al.* Multivariable characterization of atmospheric environment with data collected in flight. *Atmosphere* **13**, 1715 (2022).
99. Belacel, N., Wang, C. & Cupelovic-Culf, M. Clustering: unsupervised learning in large biological data. *Statistical Bioinformatics: A Guide for Life and Biomedical Science Researchers* 89–127 (2010).
100. WorldBank. World bank open data. Available at https://data.worldbank.org/indicator/SH.MED.PHYS.ZS?most_recent_value_desc=true (July 20,2020) (2020).
101. Statista. The countries with the most critical care beds per capita. Available at <https://www.statista.com/chart/21105/number-of-critical-care-beds-per-100000-inhabitants/> (July 20,2020) (2020).
102. Soriano, J. B. *et al.* Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The Lancet Resp. Med.* **5**, 691–706 (2017).
103. WorldPopulationReview. Global health observatory (gho) data. Available at <https://worldpopulationreview.com/> (June 6,2020) (2020).
104. EuropeanHeartNetwork. Fighting heart diseases and stroke. Retrieved from <http://www.ehnheart.org/> (2020).
105. Bikbov, B. *et al.* Global, regional, and national burden of chronic kidney disease, 1990–2017: A systematic analysis for the global burden of disease study 2017. *Lancet* **395**, 709–733. [https://doi.org/10.1016/s0140-6736\(20\)30045-3](https://doi.org/10.1016/s0140-6736(20)30045-3) (2020).
106. WorldBank. World bank open data. Available at <https://data.worldbank.org/indicator/AG.LND.TOTL.UR.K2> (July 11,2020/)(2020).
107. OWID. Statistics and research: Coronavirus pandemic (covid-19). Available at <https://ourworldindata.org/coronavirus> (July 6,2020) (2020).
108. Camacho, S. *et al.* How much for a broken heart? costs of cardiovascular disease in colombia using a person-based approach. *PLOS ONE* **13**, e0208513. <https://doi.org/10.1371/journal.pone.0208513> (2018).
109. Iso, H. A japanese health success story: Trends in cardiovascular diseases, their risk factors, and the contribution of public health and personalized approaches. *The EPMA J.* **2**, 49–57. <https://doi.org/10.1007/s13167-011-0067-0> (2011).
110. Icaza, G. *et al.* Estimation of coronary heart disease risk in chilean subjects based on adapted framingham equations. *Revista medica de Chile* **137**, 1273–1282 (2009).
111. Lab, C. D. World covid-19 daily cases with basemap. In *UNF:6:DfkQNd3HHOdRRegGu3bjgw==*, <https://doi.org/10.7910/DVN/L2LOLOT> (Harvard Dataverse, 2020).
112. WHO. Global research on coronavirus disease (covid-19). *World Health Organization* (2020).
113. UNESCO. Demographic and socio-economic. Available at http://data.uis.unesco.org/Index.aspx?DataSetCode=DEMO_DS (July 6,2020) (2020).

Author contributions

All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. They are all involved in writing the paper. A.G. is the lead author and principal investigator behind this study. He was responsible for developing the original idea and the bi-criteria qualitative method to evaluate national healthcare systems during the COVID-19 pandemic. N.B. is a leading expert in fuzzy clustering algorithms, and he played a crucial role in the project, providing his expertise in this area. He made significant contributions to every aspect of the paper, including the analysis of the results and the interpretation of the findings. L.B. is the lead expert on data envelopment analysis, made important contributions to the project and the paper. L.B. contributed to every section of the paper, ensuring its quality and accuracy. B.M. is the advanced computing specialist in the team, and he played a key role in the implementation of the algorithm and the experimental platform. M.E. is a medical doctor and clinician expert, she was responsible for the review of all population vulnerability factors to COVID-19, as well as the collection of the population health data. She participated in the literature review and assisted in the analysis of the results. H.A., a graduate student, played a supportive role in the project, helping with the data collection, cleaning, and validation. He participated in the literature review and assisted in the analysis of the results.

Funding

This research is funded by the Canadian Institutes of Health Research (CIHR grant Ref.# VR5172669) under the COVID-19 Rapid Research FO - Social Policy and Public Health Responses grant and by National Research Council Canada funding (operating grant CDTs-101-1). It was also enabled in part by the support and the computational resources provided by WestGrid (www.westgrid.ca) and Compute Canada (www.computeCanada.ca) under the COVID-19 resource allocation track on Arbutus cloud (https://docs.computeCanada.ca/wiki/Cloud_resources).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-69212-x>.

Correspondence and requests for materials should be addressed to A.G. or N.B.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© His Majesty the King in Right of Canada, as represented by the National Research Council of Canada, and Adel Guitouni, Loubna Benabbou, Belaid Moa, Munire Erman, Halim Abdul 2024