How Do Health System Features Influence Health System Performance?

How Do Health System Features Influence Health System Performance?

This work is published under the responsibility of the Secretary-General of the OECD and the Health Foundation. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD or the Health Foundation.

The names and representation of countries and territories used in this joint publication follow the practice of the OECD.

This document, as well as any data and map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Specific territorial disclaimers applicable to the OECD:

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Please cite this publication as:
OECD/The Health Foundation (2025), *How Do Health System Features Influence Health System Performance?*, OECD Publishing, Paris, https://doi.org/10.1787/7b877762-en.

ISBN 978-92-64-82221-4 (print)
ISBN 978-92-64-81481-3 (PDF)
ISBN 978-92-64-31401-6 (HTML)

Photo credits: Cover © Anchiy/Getty Images.

Corrigenda to OECD publications may be found at: https://www.oecd.org/en/publications/support/corrigenda.html.

Foreword

Health systems everywhere face major challenges – demographic pressures, funding constraints, workforce shortages, rising inequalities, increasing patients' expectations to name a few. In this context, international comparisons can serve as a tool to help policy makers identify policy options to better address these challenges. These comparisons can raise awareness of health systems' relative strengths and shortcomings, as well as stimulate policy debates that aim to improve health system outcomes.

Health systems differ in all aspects of their design: how they are governed, how they are funded, how they generate and deploy resources, and how they deliver services. While there is widespread agreement that the design of these functions influence health outcomes, limited work has been able to identify how, and to what degree, they do.

This report has two main objectives: to identify groups of countries with similar health system characteristics (health system clusters); and to compare and assess performance between and within health system clusters to assess potential links between system characteristics and performance. This work follows and builds on earlier OECD analysis of health system performance carried out in 2009-10. It draws heavily on several rounds of the OECD survey on health system characteristics carried out by the OECD Health Committee.

The analysis highlights that health systems with a different mix of institutions characteristics can achieve similar levels of efficiency. In other words, there is no best health system institutional set up that will "automatically" result in higher performance. Rather, health systems with similar institutional arrangements can learn from each other about how specific policy actions can be leveraged to improve health system performance.

Acknowledgements

The preparation and writing of this report were co-ordinated by Luca Lorenzoni. Co-authors include Aikaterini Giannadou, Nicolás Larrain and Giulia Perego of the OECD. Ruth Thorlby and Hugh Alderwick (The Health Foundation, United Kingdom) provided guidance and advice on the analysis and reviewed several versions of the report.

The authors would like to extend their gratitude to the members of the Technical Advisory Group who advised on the framework, methodologies and analytical choices made in conducting the project: Reinhard Busse, Technische Universität Berlin (Germany); Jonathan Cylus, World Health Organization Barcelona Office for Health Systems Financing, Barcelona (Spain) and European Observatory on Health Systems and Policies, London (United Kingdom); Sophie Gerkens and Nicolas Bouckaert, Belgian Health Care Knowledge Centre (KCE), Bruxelles (Belgium); Philip Haywood, University of Sydney, Sydney (Australia) – the Chair of the group; Irene Papanicolas, Brown University, Providence (United States); Veena Raleigh, The King’s Fund, London (United Kingdom); Marco Varkevisser, Erasmus University, Rotterdam (the Netherlands); and Claus Wendt, University of Siegen, Siegen (Germany). Those experts did not represent countries and served in an honorary capacity. This report would not have been possible without their contributions and support.

The authors would also like to extend their gratitude to the participants in the OECD Health Committee who reviewed and commented on previous versions of the report, and to the many government officials and policy and technical experts across the OECD who responded to several rounds of the OECD health systems characteristics survey. Neither they nor their institutions are responsible for any of the opinions expressed in this report.

At the OECD, the authors wish to thank Francesca Colombo, Frederico Guanais and Mark Pearson, from the Directorate of Employment, Labour and Social Affairs, for their comments and suggestions. Thanks also go to Isabelle Vallard for her valuable administrative support, and to Lucy Hulett for editorial and media assistance.

The work has received financial support from The Health Foundation.

Table of contents

FIGURES

TABLES

Executive summary

International comparisons are an important tool for assessing health system performance and can raise awareness of health systems' relative strengths and shortcomings, facilitating international learning and stimulating policy debates. Cross-country comparisons are widely used and valued by decision makers, but the most useful comparisons may not be the most obvious ones. The selection of the most policy relevant "comparator" countries should be determined by the policy question at hand. For example, if a Swiss policy maker were looking for examples on how to improve their primary care system, looking at the British experience may be less useful than looking at countries such as the Netherlands, where the healthcare structure has some similarities to their own. This may mean using different categorisations of health systems based on the characteristics that are most relevant for each policy question.

Clustering is a technique which can be used to form groups of similar health systems that share distinct properties. These shared characteristics might not be visible by simply exploring distributions and studying the effect of system features directly on the outcome of choice. To make valid and useful comparisons of performance, health systems can be clustered based on key qualitative features that underpin healthcare funding and delivery in a given country, for example arrangements to organise population coverage, the financing of healthcare insurance and delivery, the organisation of healthcare delivery, and key aspects of governance and resource allocation.

The analyses contained in this report considered whether different group of counties sharing similar health system characteristics appear to perform better relative to others. Specifically, a focus is whether health system efficiency – a measure that offers insights into performance – is influenced by the overall design of health systems. Health systems of different OECD countries have been clustered based on indicators constructed from countries' responses to the OECD survey on health systems characteristics. Indicators considered in defining the clusters are the degree of user choice of basic health coverage, degree of public or private provision of primary care and outpatient specialist services, the degree of patient choice of providers, health insurance as a secondary source of coverage ("over the basic" coverage), and the role of primary care in the health system (gate-keeping). Those indicators were selected as those that most differentiate health systems, and they are the same characteristics that were used to cluster OECD health systems in a previous analysis undertaken in 2008-10. Nevertheless, there is an element of judgement that is involved in choosing these indicators, rather than others. An efficiency measure was then derived using health spending as a share of GDP as input and age standardised mortality rate as output.

The analysis confirms the results of previous OECD work suggesting that there is no single health system design that is most associated with higher efficiency. In other words, there is no indication that any one group of health systems would systematically outperform another.

The analysis suggests that there is room for health systems sharing the same broad characteristics to improve performance by adopting policy actions to improved efficiency that might borrow elements from other systems. Rather than engaging in wholesale reforms, improving health system efficiency requires the use of more targeted policies to respond to a specific policy question.

To this aim, two additional sets of analysis look at some actionable policy levers available to countries across clusters that seem to be particularly promising for countries to improve performance, regardless of their institutional set-up.

A second set of analyses looks at the role of provider payment mechanisms. The analysis considers whether differences in treatable mortality rates across countries can be explained by clustering health systems based on volume incentives embedded in physicians' payment schemes and financial incentives for healthcare quality. A third set of analyses then looks at the role of primary healthcare. The analysis considers whether differences in avoidable hospital admissions for selected conditions – asthma, chronic obstructive pulmonary disease and congestive heart failure – can be linked to health systems being more primary care oriented.

Results show that health systems providing a higher degree of incentives for quality to providers may also achieve better access to high-quality care, helping to reduce treatable mortality rates, as compared to health systems relying on limited incentives for quality and more traditional fee-for-service payments.

Results also indicate that primary care oriented health systems – defined by a stronger role of General Practitioners (GPs) as gatekeepers, better care continuity and stronger financial incentives directed to primary care physicians to improve quality – also display lower avoidable hospitalisations, a variable often used in the relevant literature as a marker of quality and access to primary care.

It is important to note that the relationship between health system clusters and outcome variables should not be interpreted as causal, as the methodology developed to answer the above questions is underpinned by several limitations. In reality, health systems are more nuanced than described in a set of health system indicators. Boundaries between different groups of health policies and institutions are rarely clear-cut. Thus, the panel dataset of health system features provides as complete a picture as possible of the contextual variation across OECD health systems, given data availability. Furthermore, findings are limited by the outcome variables available. In addition, cluster analysis algorithms will always produce a result (a set of clusters) whether there are true patterns in the data or not. When there are natural clusters in the data, there is no way of knowing whether the algorithm has clustered the data "correctly" as there are no "right answers" with which to compare. Finally, the effects of clusters on the outcome variables may be due to health systems features, lifestyle or socio-economic factors that are not fully captured or controlled for in the analyses.

Despite these limitations, this type of analysis has a real value added for policy makers seeking to understand how changes in certain system features might influence performance. This analysis is also the best of the kind that could be done given the data available. Furthermore, while this report considered two set analyses around how actionable policy levers influence performance, the same analytical approach could be used to understand the impact of many other policy levers, using the rich data set of health systems characteristics.

1 Opportunities and challenges for the comparison of health system performance based on their institutional characteristics

The chapter discusses the importance and challenges of comparing health system performance across countries. While international comparisons are valuable for identifying strengths and weaknesses in healthcare systems, making improvements based on these comparisons is complex due to variations in governance, funding, resources, and service delivery across different countries. The chapter focuses on cluster analysis as a methodological approach to grouping and comparing health systems and discusses two main approaches to clustering: grouping health systems based on their overall characteristics, such as financing methods, or type of coverage; targeted policy-based clustering, with a focus on specific policy questions to help make actionable policy improvements.

There is a long history of comparing health system performance. International comparisons are recognised as being an important tool for assessing performance and prompting improvement (Papanicolas and Smith, 2013[1]; OECD, 2024[2]) as they can raise awareness of health systems' relative strengths and shortcomings, facilitating international learning and stimulating needed policy debates.

Yet determining how to improve health system performance based on international comparisons is fraught with challenges (Bowden, Figueroa and Papanicolas, 2024[3]). Health systems can differ in many ways, for example in how they are governed, how they are funded, how they generate and deploy resources, and how they deliver services. While there is widespread agreement that these features influence health system performance, it is more difficult to assess how much they matter, which ones matter most, how they individually affect different dimensions of performance, and how they are affected by the wider context in which they operate – be that the rest of the health system, other country-specific factors or the social determinants of health.

Data on factors that affect performance – such as health system features, the demographic characteristics of the populations being compared and the broader political, socio-economic and cultural context within countries – can be used to make more meaningful empirical comparisons through using them to identify comparators (groups or clusters of health systems), to adjust for exogenous variation or to help interpret results (Jacobs, Smith and Street, 2006[4]; Papanicolas and Marino, 2024[5]).

Cluster analysis (James et al., 2021[6]; Hastie, Tibshirani and Friedman, 2009[7]) is a useful descriptive tool that can be used to group health systems as it provides valuable insights into a dataset where there are distinct populations, which might not be seen by simply exploring distributions and comparing data parameters. For example, it has been recently used by OECD to assess the transferability of public health interventions (Wiper et al., 2022[8]).

Various approaches have been developed to cluster health systems in the past (Ferreira et al., 2018[9]). These generally use range of data on the characteristics of health systems (and sometimes their capacity, performance, and other factors) to identify typologies (clusters) based on the overall design or key features of health systems. This can include data on more or less modifiable characteristics – e.g. revenue sources versus payment systems – depending on the approach and aims of the researchers. As an example, Reibling et al. (2019[10]), with the objective of identifying groups or clusters of peers, developed a typology of five healthcare systems based on supply, public/private mix, access regulation, primary care orientation and performance. Ferreira et al. (2018[9]) used three factors – an aggregate of health systems financing, medical doctors per 100 000 population and hospital discharges due to diabetes, hypertension or asthma per 100 000 population – to identify five clusters of countries. Gabani et al. (2023[11]) used health expenditures by financing scheme as a share of total health expenditure to identify three clusters of countries – those where health expenditure is channelled predominantly via a government-funded arrangement, a contributory social health insurance arrangement or an out-of-pocket arrangement. Paris et al. (2016[12]) use the type of primary coverage to identify four clusters of health systems: residence-based; contributory, single payer; contributory, multiple insurers with automatic affiliation; contributory, multiple insurers with choice of insurer.

Another approach is to cluster health systems based on more targeted policy questions (Papanicolas et al., 2024[13]). This could include how countries use a particular policy lever (e.g. payment systems) or how they approach a particular policy problem (e.g. regulating new technology). This approach has the potential to offer more actionable insights for policy improvement.

Both of these approaches to clustering are valid depending on the question being asked and the intended use of the comparison. But more targeted approaches are underused and could be developed using new, richer, data such as the OECD's Health System Characteristics (HSC) survey (see Annex A), the European Observatory's Health System in Transition profiles and the WHO Health Financing Progress Matrix.

The aim of this work is to understand the links between health systems overall design and performance of OECD countries on key health system indicators. New OECD data on health system characteristics were used to develop updated clusters of health systems based on their overall design and policy approach. Health system performance was compared between and within these clusters to understand potential links between overall health system policies and efficiency. A more targeted approach was then taken to clustering and performance comparison based on selected policy questions – including how payment systems might shape performance, and differences in performance between more or less primary care oriented systems – to help identify actionable policy levers that – independently of the overall design of a health system – could improve performance.

References

Bowden, N., J. Figueroa and I. Papanicolas (2024), "Bridging borders: current trends and future directions in comparative health systems research", *Health Services Research*, https://doi.org/10.1111/1475-6773.14385. [3]

Ferreira, P. et al. (2018), "EU health systems classification: a new proposal from EURO-HEALTHY", *BMC Health Services Research*, Vol. 18/1, https://doi.org/10.1186/s12913-018-3323-3. [9]

Gabani, J., S. Mazumdar and M. Suhrcke (2023), "The effect of health financing systems on health system outcomes: A cross-country panel analysis", *Health Economics*, Vol. 32/3, pp. 574-619, https://doi.org/10.1002/hec.4635. [11]

Hastie, T., R. Tibshirani and J. Friedman (2009), *The Elements of Statistical Learning*, Springer New York, New York, NY, https://doi.org/10.1007/978-0-387-84858-7. [7]

Jacobs, R., P. Smith and A. Street (2006), *Measuring efficiency in health care*, Cambridge University Press. [4]

James, G. et al. (2021), *An Introduction to Statistical Learning*, Springer US, New York, NY, https://doi.org/10.1007/978-1-0716-1418-1. [6]

OECD (2024), *Rethinking Health System Performance Assessment: A Renewed Framework*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/107182c8-en. [2]

Papanicolas, I. et al. (2024), "Policy questions as a guide for health systems' performance comparisons", *Bull World Health Organ*, pp. 550-552, https://doi.org/10.2471/BLT.24.291635. [13]

Papanicolas, I. and A. Marino (2024), *International comparisons: who has the best health system in the world?*, Edward Elgar Publishing. [5]

Papanicolas, I. and P. Smith (2013), *Health system performance comparison: an agenda for policy, information and research*, Open University Press. [1]

Paris, V. et al. (2016), "Health care coverage in OECD countries in 2012", *OECD Health Working Papers*, No. 88, OECD Publishing, Paris, https://doi.org/10.1787/5jlz3kbf7pzv-en. [12]

Reibling, N., M. Ariaans and C. Wendt (2019), "Worlds of Healthcare: A Healthcare System Typology of OECD Countries", *Health Policy*, Vol. 123/7, pp. 611-620, https://doi.org/10.1016/j.healthpol.2019.05.001. [10]

Wiper, O. et al. (2022), “Cluster analysis to assess the transferability of public health interventions”*, OECD Health Working Papers*, No. 133, OECD Publishing, Paris, https://doi.org/10.1787/a5b1dcc1-en. [8]

2 Does one "best" health system exist?

The chapter updates previous OECD work on health system efficiency, using newer data on health systems characteristics and an expanded set of countries. The analysis used cluster analysis to group countries with similar healthcare system characteristics, resulting in eight distinct clusters based on features like market mechanisms, provider choice, insurance coverage, and gatekeeping arrangements. There is no evidence that any particular type of health system consistently outperforms others. High-performing countries can be found in all clusters and health systems doing poorly are also present in all groups. The analysis concludes that rather than attempting wholesale system changes, countries should focus on implementing specific policy improvements that can enhance performance regardless of their overall healthcare system design.

This chapter looks at whether efficiency – a measure that offers insights into performance – is influenced by the overall design of health systems. New OECD data on health system characteristics were used to understand links between health policy and performance.

Previous OECD work (Joumard, André and Nicq, 2010[1]) assessed whether higher efficiency is associated with the overall design and policy approach of health systems. Those empirical analyses (Box 2.1) suggested that there is room in all health systems to improve performance, and there is no healthcare system that performs systematically better in delivering cost-effective healthcare. Those analyses also showed that increasing the coherence of policy settings, by adopting best policy practices within a similar system and borrowing the most appropriate elements from other systems will likely be more practical and effective to raise healthcare spending efficiency – big bang reforms[1] are therefore unlikely to be warranted.

Box 2.1. Results of previous OECD work on health systems efficiency

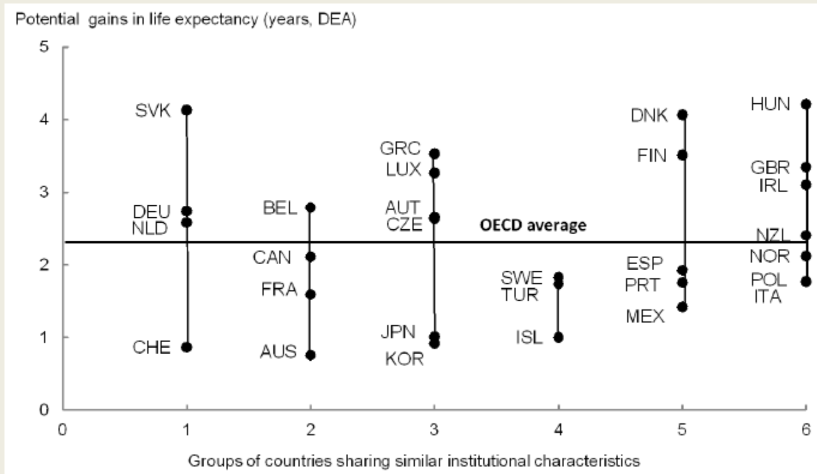
To capture the key features of the overall design and policy approach of health systems, the following indicators – constructed on the basis of responses to the 2008 round of the OECD survey on health systems characteristics – were used to identify clusters of health systems:

Indicator	Description
Degree of user choice of basic coverage	Evaluates the source of basic healthcare coverage and whether there is ability/freedom to choose an insurer. residence-based; insurance-based, single insurer; insurance-based, multiple insurers without choice; insurance-based, multiple insurers with choice
Degree of private provision of primary care and outpatient specialist services	Evaluates the degree of private provision in primary care and outpatient specialist care. A higher value of the indicators means that the predominant provision of primary and outpatient specialist care is private
Patient choice of providers	Indicates whether individuals are free to choose any doctor or hospital, face incentives to choose any specific doctor or hospital, or have a limited choice. A higher value of the indicator means a larger possibility to use any provider
Health insurance as a secondary source of coverage ("over the basic" coverage)	Shows spending on private voluntary health insurance as a secondary source of coverage as a share of total health expenditure A higher value of the indicator means a larger spending on secondary coverage
Role of primary care in the health system (gate-keeping)	Indicates whether the referral from a primary care physician to access specialist care is required or there are financial incentives to do so. A higher value of the indicator means that referral is required

Those indicators were selected as those that most differentiate health systems – using a Principal Component Analysis[1] – and that help identify plausible and interpretable clusters (Joumard, André and Nicq, 2010[1]).

The empirical analyses suggested that efficiency levels vary more within groups of countries sharing similar institutional characteristics than between groups. Thus, there is no indication that one group of healthcare system would systematically outperform another. On the contrary, countries performing well can be found in all institutional groups. Countries doing poorly are also present in most groups.

Figure 2.1. Efficiency scores across and within country groups for 2007



Note: potential gains in life expectancy are derived from an output-oriented DEA with per capita health spending and a composite indicator of socio-economic environment and lifestyle factors as inputs.

Source: Joumard, I., C. André and C. Nicq (2010[1]), "Health Care Systems: Efficiency and Institutions", https://doi.org/10.1787/5kmfp51f5f9t-en.

1. Principal Component Analysis is a multivariate statistical method that combines information from several variables observed on the same subjects into fewer variables, called principal components (Hotelling, 1933[2]).

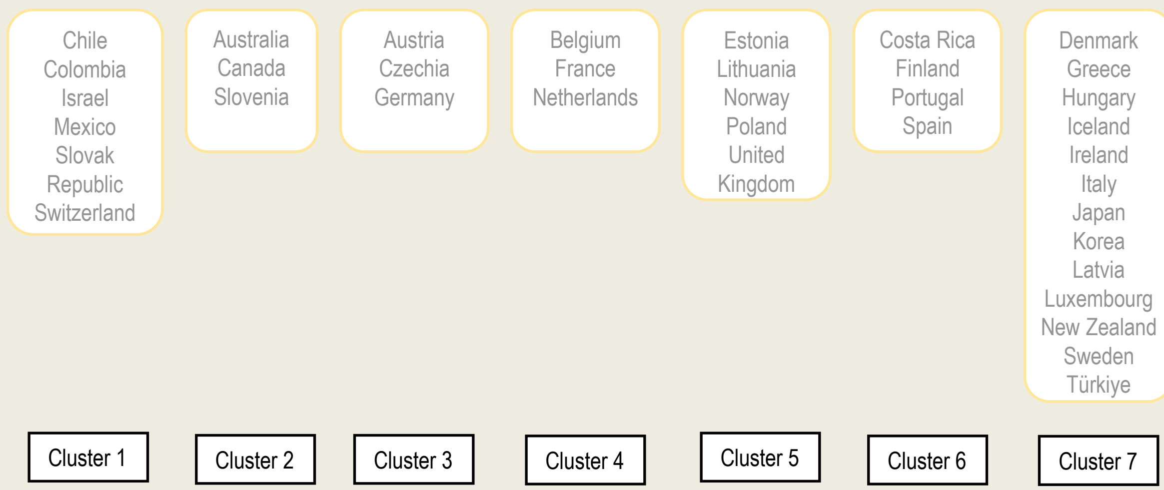
This chapter updates those analyses to see whether findings are confirmed using a larger set of countries, responses to the 2016 and 2023 rounds of the HSC survey,[2] and different input and output variables.

First, a pure statistical data driven approach to clustering was conducted based on the five indicators used in previous OECD work – that is degree of user choice of basic coverage; degree of private provision of primary care and outpatient specialist services; patient choice of providers; health insurance as a secondary source of coverage ("over the basic" coverage); role of primary care in the health system (gate-keeping) (Box 2.2).

Box 2.2. Results of the data driven approach to clustering health systems

The data driven approach identified the following seven clusters of health systems:

Figure 2.2. Clusters of health systems



At first sight, some of the clusters look plausible – most of them include countries which are neighbours of each other and/or which have aspects in common. However, there are interlopers that are difficult to explain (e.g. Slovenia in cluster 2; Costa Rica in cluster 6), and cluster 7 groups quite different health systems.

Cluster analysis algorithms will always produce a set of clusters whether there are true patterns in the data or not; however, it is important that results are plausible, interpretable, and easily explained to an audience that is not necessarily familiar with statistical methods. Therefore, selecting the appropriate number of homogeneous groups or families of health systems for policy discussions should not be based on quantitative measures alone as the challenge lies in finding the optimal balance to ensure that comparisons are both valid and policy relevant.

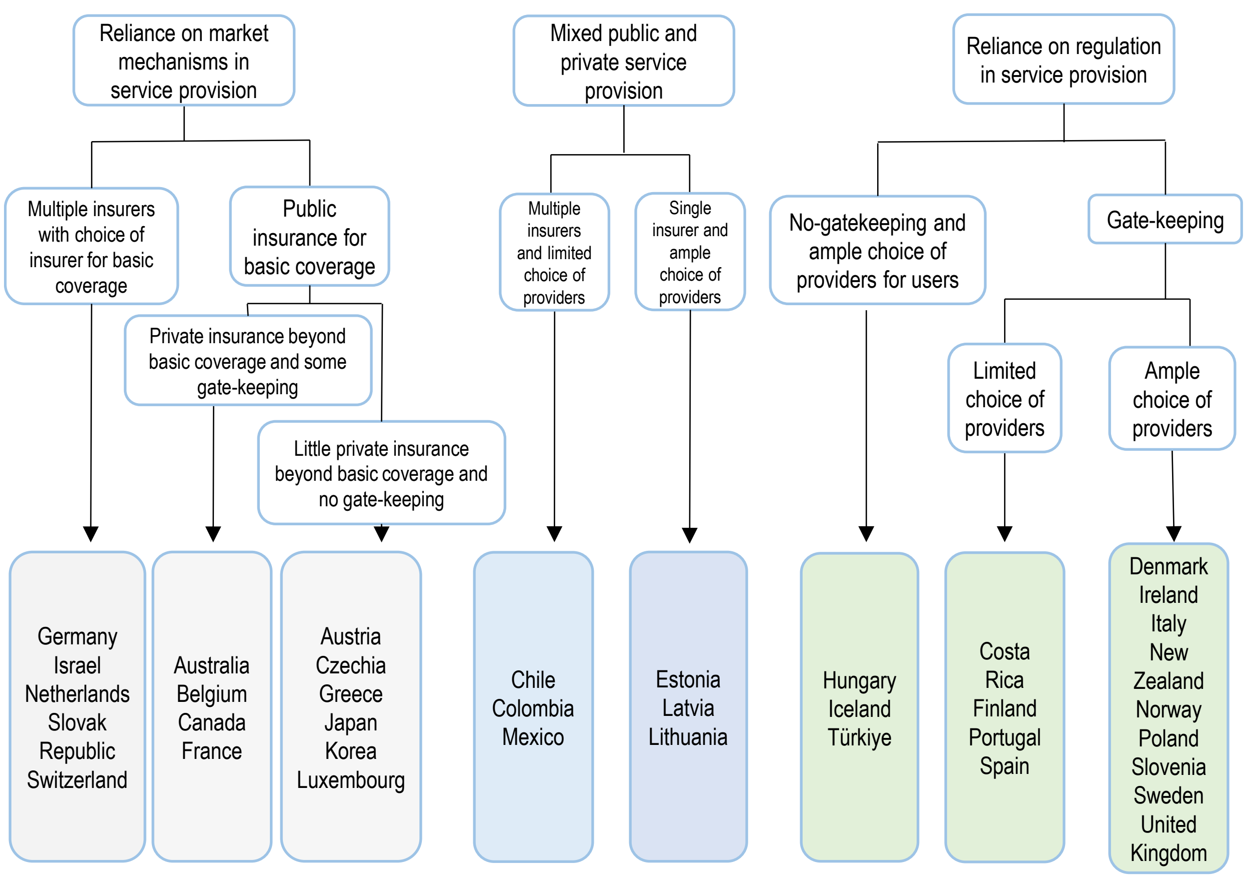
Based on the data driven approach, some expert judgement was then used to ensure that there are meaningful and identifiable policy differences to explain why countries are grouped together. As a result, eight clusters that display the following features were identified (Figure 2.3):

- Cluster 1: Germany, Israel, the Netherlands, the Slovak Republic and Switzerland rely extensively on market mechanisms to steer the behaviour of service providers, and basic insurance coverage is provided by multiple insurers with choice of insurer on the part of users.
- Cluster 2: A second group of countries – Australia, Belgium, Canada and France – relies extensively on market mechanisms in service provision and features public basic insurance coverage. Over-the-basic insurance coverage plays a significant role, and cost control generally takes the form of moderate gate-keeping arrangements.
- Cluster 3: The third group – which includes Austria, Czechia, Greece, Japan, Korea and Luxembourg – is also characterised by extensive use of market mechanisms in private provision of care. But there is no gate-keeping system in place and over-the-basic coverage is limited.
- Cluster 4: Chile, Colombia and Mexico feature a mixed provision of healthcare services, basic coverage provided by multiple insurers and limited patient choice of providers.

- Cluster 5: A fifth group of countries – Estonia, Latvia and Lithuania – is characterised by a mixed provision of healthcare services, basic coverage provided by a single insurer and ample patient choice of providers.
- Cluster 6: The healthcare systems of Hungary, Iceland and Türkiye offer free choice of provider to patients in all three areas of care – primary, specialist and hospital care – with no gate-keeping.
- Cluster 7: In the group consisting of Costa Rica, Finland, Portugal and Spain, healthcare is mainly provided by a heavily regulated public system. Patients' choice among providers is very limited and the role of gate-keeping is important.
- Cluster 8: The last group also consists of heavily regulated public systems – Denmark, Ireland, Italy, New Zealand, Norway, Poland, Slovenia, Sweden and the United Kingdom. Compared with the previous group, the possibility for patients of choosing between providers tends to be large.

As compared to the previous OECD analysis (Joumard, André and Nicq, 2010[1]), two clusters were added to group health systems with mixed private-public provision (Cluster 4 and Cluster 5). The other clusters (1-3, 6-8) do not look substantially different from earlier analysis, reflecting that not many countries saw major changes in their health system characteristics. Only Mexico was shifted to the new Cluster 4 as its health system presents a mixed provision of healthcare services, Hungary to Cluster 6 as there is no referral required to access specialist care (gate-keeping) and Sweden to Cluster 8 as referral is required to access specialist care.

Figure 2.3. Groups of health systems with similar institutional features



Statistical methods

Data Envelopment Analysis (DEA) – a nonparametric statistical technique[3] – is a commonly used analytical tool to estimate the relative efficiency with which inputs are turned into output (Jacobs, Smith and Street, 2006[3]). DEA was used in the comparison of efficiency of OECD health systems mentioned earlier (Joumard, André and Nicq, 2010[1]). It was also used in other health system efficiency analyses (Sicari and Sutherland, 2023[4]; Moran, Suhrcke and Nolte, 2023[5]).

Health spending as a share of GDP[4] was the main input variable in the DEA model used in this set of analyses. The variable proxies the countries' current investments aiming to improve population health, while accounting for the different levels of GDP across OECD countries.

The selection of output variables is another important methodological decision to be made, as it should capture key results of health systems that will reflect their efficiency. Depending on the purposes of the analysis, these could be, for example, health outcomes, metrics associated with service delivery, indicators of quality of care, indicators of access to care, or even indicators related to carbon emissions.

A wide range of health outcome variables have been previously used to compare change over time and between populations, ranging from life expectancy (Joumard, André and Nicq, 2010[1]; Medeiros and Schwierz, 2015[6]) to infant mortality (Retzlaff-Roberts, Chang and Rubin, 2004[7]; Manavgat and Audibert, 2024[8]) and treatable mortality (Medeiros and Schwierz, 2015[6]).

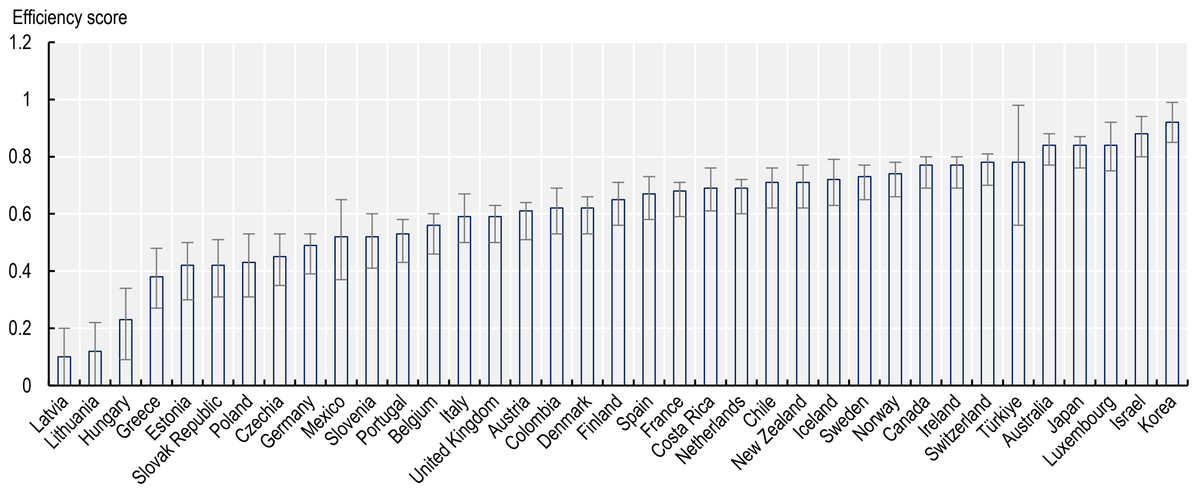
As comparisons of measures such as life expectancy at birth do not take into account differences in population structures across countries, in contrast to previous OECD analyses, this set of analyses uses yearly age standardised mortality rates (ASMR)[5] as the output variable in the DEA model. Age standardised measures account for differences and changes in population structure and size that can impact mortality-based health outcomes across countries and in a given country over time. For example, if the overall population is ageing over time, then it might be reasonably expected that the observed rates of deaths would also increase, independently of the performance of the health system. Furthermore, ASMR does not take into account how other external factors – including environmental and social factors and health behaviours – that influence performance.

A DEA output-oriented efficiency measure is then compared across countries to explore the proportional expansion in output that are possible both within and between groups/clusters of health systems.

Key findings

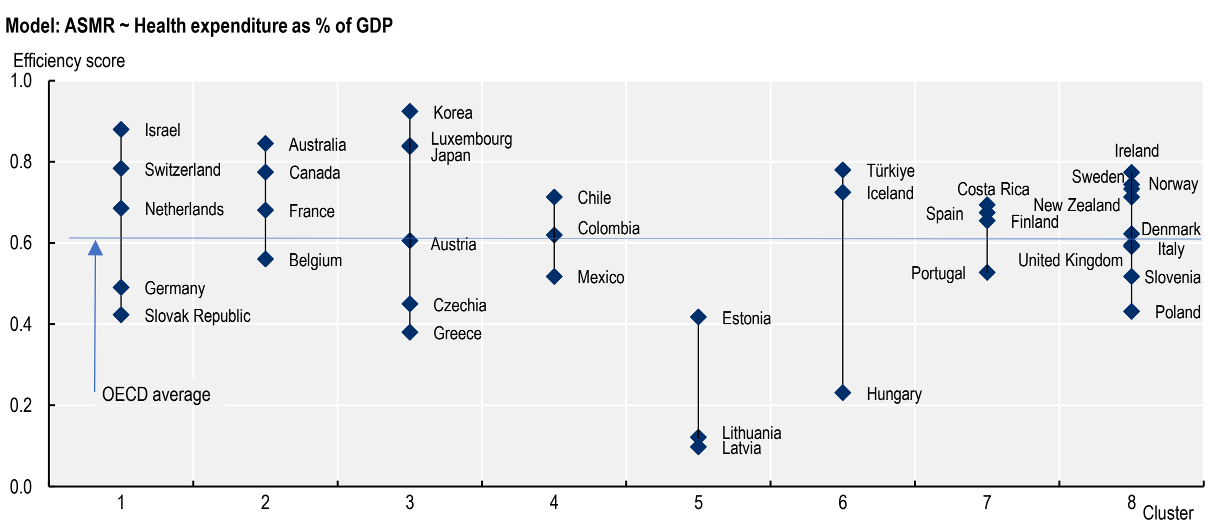
Efficiency scores by country computed using DEA are shown in Figure 2.4.

Figure 2.4. Efficiency scores with confidence intervals by country, 2019



Descriptive analysis showed that health systems performing well can be found in all institutional groups. Health systems doing poorly are also present in all groups, in particular in cluster 5 characterised by a mixed provision of healthcare services, basic coverage provided by a single insurer and ample patient choice of providers (Figure 2.5). This indicates that no institutional characteristics are suggestive of "best performance", but there is ample possibility for within and across groups learning.

Figure 2.5. Efficiency scores within and across health systems clusters, 2019



Note: Efficiency score 1 = highest efficiency.

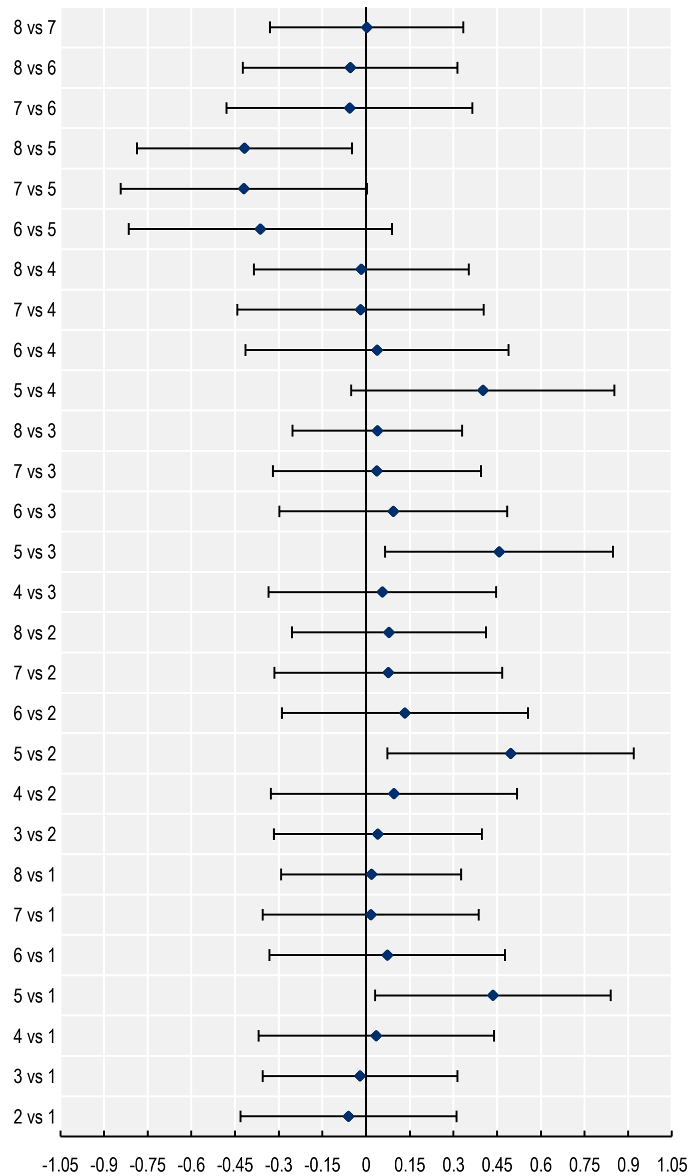
Analysis of variance (ANOVA) indicated that more than almost 40% of the variation observed in efficiency scores by countries is explained by the eight clusters used in this analysis (Table 2.1).

Table 2.1. Average efficiency scores and confidence intervals by health systems clusters, 2019

Cluster	Countries	Efficiency score (1 = highest efficiency)	
		Mean	95% confidence interval
1	Germany, Israel, Netherlands, Slovak Republic, Switzerland	0.65	0.48 – 0.82
2	Australia, Belgium, Canada, France	0.71	0.59 – 0.84
3	Austria, Czechia, Greece, Japan, Korea, Luxembourg	0.67	0.49 – 0.85
4	Chile, Colombia, Mexico	0.62	0.51 – 0.73
5	Estonia, Latvia, Lithuania	0.21	0.01 – 0.41
6	Hungary, Iceland, Türkiye	0.58	0.24 – 0.92
7	Costa Rica, Finland, Portugal, Spain	0.64	0.56 – 0.71
8	Denmark, Ireland, Italy, New Zealand, Norway, Poland, Slovenia, Sweden, United Kingdom	0.64	0.56 – 0.71
Total		0.61	0.55 – 0.67
Variation (η^2)			
	Between groups	39.8%	
	Within groups	60.2%	

Furthermore, the difference in the mean efficiency score by pairs of clusters was assessed using the Tukey 95% family-wise confidence levels. Even though the difference in the mean score (the point) looks as if one cluster is better (more efficient) than another, there is enough variation for us not be 95% sure that this is the case. Only the mean efficiency score of cluster 5 is significantly lower than the mean efficiency score of clusters 1, 2, 3 and 8 (Figure 2.6).

Figure 2.6. Comparison of the difference in the mean efficiency scores for all pairs of clusters



Note: For each pair of health system clusters, the difference in the means (the blue dot) and its confidence interval are shown.

In this analysis, some countries with low levels of health spending as a share of GDP appear high on the efficiency score (e.g. Türkiye), in contrast to some historically well-funded health systems (e.g. Germany). This result may be attributed to the declining marginal productivity observed in countries with high health expenditure as a share of GDP. In support of these long-standing OECD findings (OECD, 2017[9]), the International Monetary Fund recently showed that the relationship between financial resources and health outcomes is “flatter” for advanced economies as compared to emerging markets and developing countries, thereby additional health spending results in smaller gains in health outputs over time (Garcia-Escribano, Mogues and Juarros, 2022[10]). Taken together, these results support the idea that as inputs increase in health systems of major developed economies, outputs tend to increase too, but at a slower rate (Gallet and Doucouliagos, 2017[11]). This helps explain the lower efficiency of health systems of countries with higher level of inputs.

Sensitivity analyses

Sensitivity analyses were conducted to assess the robustness of results to adding one input variable to the base model, using efficiency scores estimated using DEA for each year between 2016-22 as an input in a regression analysis, and using life expectancy at birth as outcome variable.

When the base one input – one output model is compared against different two-input models, results show that there is little variation in efficiency scores between models (Table 2.2).

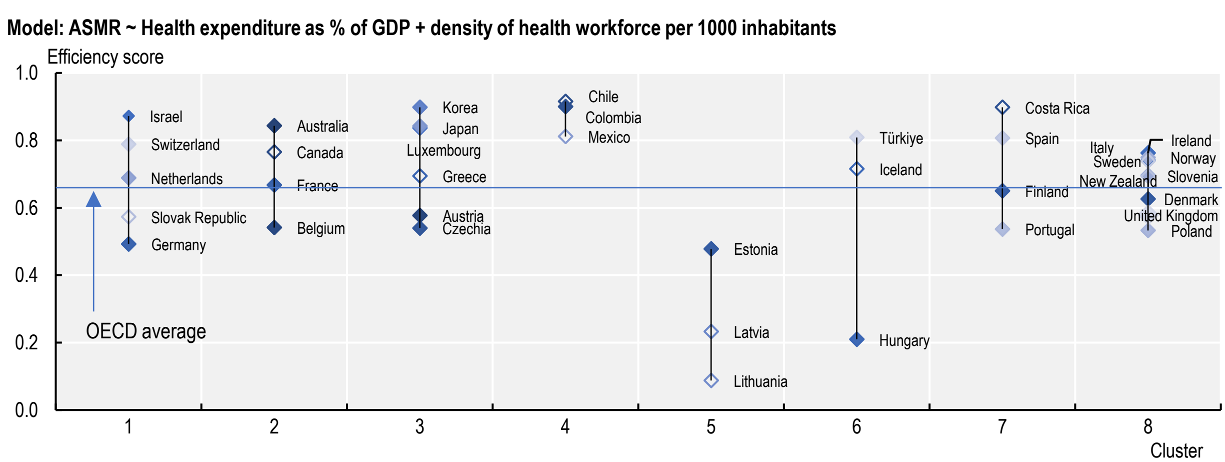
Table 2.2. Correlation of efficiency scores between models

	Base model + GDP per capita (in Purchasing Power Parities)	Base model + Obesity (% share of population)	Base model + tobacco (% share of population)	Base model + pollution	Base model + hospital beds	Base model + workforce
Correlation with efficiency scores of the base model	0.93	1	0.94	0.94	0.93	0.87

Note: Correlation is measured using the Pearson correlation coefficient.

Results also confirm the central findings that health systems performing well can be found in all institutional groups, and health systems doing poorly are also present in all groups (see Annex 2.A for details). Only when adding workforce as an input variable, results show some changes as efficiency scores of health systems with lower density of workforce – such as those assigned to cluster 4 – increase (Figure 2.7).

Figure 2.7. Efficiency scores within and across health systems clusters, 2019



1. Efficiency score 1 = highest efficiency.
2. "workforce" includes the number of persons (head counts) working in healthcare and social work (ISIC section VQ, codes 86-88).
Source: OECD Health Statistics, 2024.

To assess the robustness of results over time, efficiency scores were estimated through DEA for each year between 2016-22, Those scores were then used as input in a regression analyses. By comparison to a reference cluster, results confirm that cluster 5 shows a lower efficiency (Table 2.3) as the value of the coefficient is positive (1.09) and statistically significant with a 1% chance of this finding be wrong.

Table 2.3. Only mean efficiency score of cluster 5 is statistically different to cluster 1

Summary of beta regression results

Variable	Estimate	SE	P-values
GDP	-1.89E-05	6.74E-06	0.01**
Education	-3.18E-02	8.59E-03	0***
Gini	-5.9 267	1.4 482	0***
Post_Covid	3.19E-01	2.03E-01	0.12
Cluster 2 (versus cluster 1)	-2.35E-01	2.15E-01	0.27
Cluster 3 (versus cluster 1)	-4.26E-02	9.84E-02	0.67
Cluster 4 (versus cluster 1)	3.50E-01	3.72E-01	0.35
Cluster 5 (versus cluster 1)	1.5 779	3.57E-01	0***
Cluster 6 (versus cluster 1)	-6.71E-02	3.52E-01	0.85
Cluster 7 (versus cluster 1)	1.16E-02	1.56E-01	0.94
Cluster 8 (versus cluster 1)	-1.40E-02	1.40E-01	0.92
Pseudo-R2	1.99E+01	4.6 438	0***

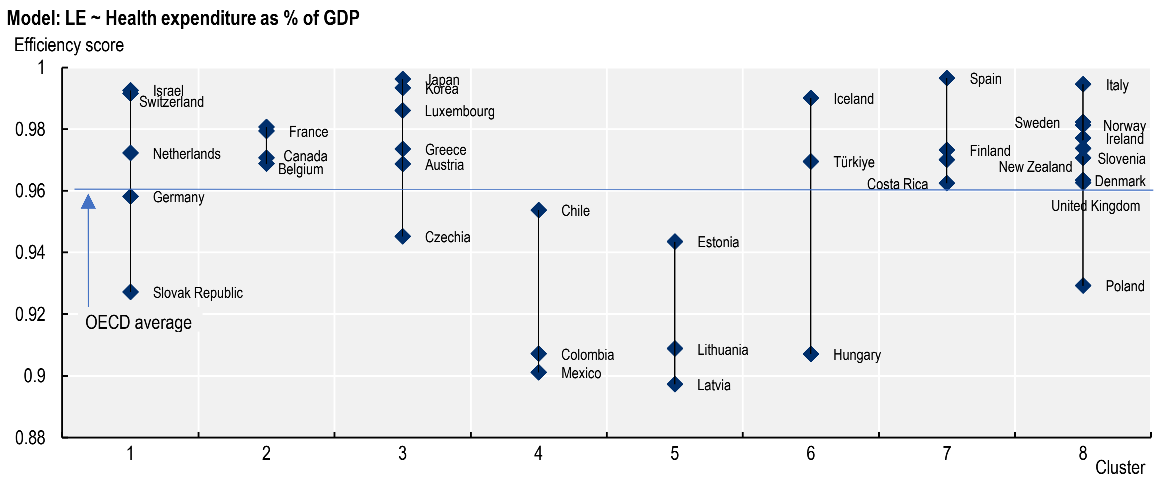
Note: Significant result at *0.05, **0.01, ***0.001 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White). The outcome of the model (Potential for outcome increase) is log transformed. The full regression results can be found in Annex 2.A.

Finally, sensitivity analyses of results to a different outcome variable were conducted. Efficiency scores are considerably higher when using life expectancy as output variable instead of age standardised mortality rates (Figure 2.8). This is explained by a lower variation in the outcome variable, making every country relatively closer to the efficiency frontier.

Compared to the base model of analysis, most countries maintain their relative position in the cluster they are assigned to, and most clusters maintain the same position relative to the OECD average. However, there are some exceptions, as life expectancy at birth does not standardise for differences in population structure. This can help explain the higher efficiency scores – compared to the base model – for Italy, Japan and Spain, countries with relatively older populations with high life expectancy, and the lower efficiency scores – compared to the base model – for countries assigned to cluster 4, countries with a relatively younger populations and lower life expectancy.

Figure 2.8. Efficiency scores within and across health systems clusters, 2019

Life expectancy at birth as output variable



Note: Efficiency score 1 = highest efficiency.

Conclusion: no best health system

In line with previous work (Joumard, André and Nicq, 2010[1]), this analysis confirmed that there is no indication that one group of health systems systematically outperform another. The policy implication of these findings is to reinforce the suggestion that large scale, "big bang" reforms, that require considerable political capital and financial resources are to be designed and implemented with caution, as just changing the whole system will not automatically improve performance. Future work could improve our ability to capture additional meaningful characteristics of health systems and explore the use of different outcome variables to refine the model for thinking about how characteristics may influence the outcomes at system level.

In addition to looking at the performance of the whole health system, there are various specific dimensions on which to evaluate health systems, and a system may perform differently on those dimensions. Therefore, there is room in health systems to adopt policy action that will lead to improved performance related to a specific dimension by understanding the most appropriate elements from other systems, regardless of their institutional set-up. To this aim, more targeted measures of performance and more granular classifications of health systems based on characteristics such as the use of financial incentives to providers to improve quality and the strength of gatekeeping (Table 2.4) may be used. Those targeted measures of performance can plausibly be more influenced by the characteristics being examined.

As those characteristics seem to be particularly promising actionable policy levers for countries to improve performance (OECD/WHO, 2014[12]; OECD, 2016[13]; OECD, 2020[14]), the following chapters help identify the role of financial incentives to providers for quality, physicians' payment methods and strength of primary care in improving performance.

Table 2.4. Actionable policies related to financial incentives to providers, physicians' payment methods and strength of gate-keeping and continuity of primary care

Cluster/country	Strength of financial incentives to providers to improve quality of care	Use of fee-for-service payments to physicians	Role of primary care in the health system (gate-keeping)	Continuity of care
Cluster 1				
Germany	Weak	Large	Medium	Strong
Israel	Weak	Limited	Medium	Strong
Netherlands	Limited	Large	Strong	Strong
Slovak Republic	Strong	Limited	Strong	n.a.
Switzerland	Weak	Large	Medium	Strong
Cluster 2				
Australia	Weak	Large	Strong	Strong
Belgium	Weak	Large	Medium	Strong
Canada	Weak	Large	Strong	Medium
France	Strong	Large	Medium	Medium
Cluster 3				
Austria	Weak	Limited	Weak	Medium
Czechia (2023)	Strong	Large	Weak	Strong
Greece	Weak	Limited	Weak	Weak
Japan	n.a.	n.a.	Weak	Weak
Korea	Strong	Large	Weak	Weak
Luxembourg	Weak	Large	Weak	Strong
Cluster 4				
Chile	Strong	Limited	Strong	Weak
Colombia	Weak	Limited	Strong	n.a.
Mexico	Weak	Limited	Strong	n.a.
Cluster 5				
Estonia	Weak	Limited	Strong	Strong
Latvia	Weak	Large	Medium	Strong
Lithuania	Weak	Limited	Strong	Strong
Cluster 6				
Hungary	Weak	Limited	Weak	Strong
Iceland	Weak	Limited	Weak	Weak
Türkiye (2016)	Weak	Large.	Weak	n.a.
Cluster 7				
Costa Rica	Weak	Limited	Strong	Strong
Finland	Weak	Limited	Strong	Strong
Portugal	Strong	Limited	Strong	Medium
Spain	Strong	Limited	Strong	Medium
Cluster 8				
Denmark	Weak	Large	Medium	n.a.
Ireland	Weak	n.a.	Strong	Medium
Italy	Weak	Limited	Strong	Medium
New Zealand	n.a.	n.a.	n.a.	n.a.
Norway	Weak	Limited	Strong	Strong
Poland	Limited	Limited	Strong	Medium
Slovenia	Weak	Limited	Strong	Strong
Sweden	Limited	Limited	Medium	Medium
United Kingdom	Strong	Limited	Strong	Strong

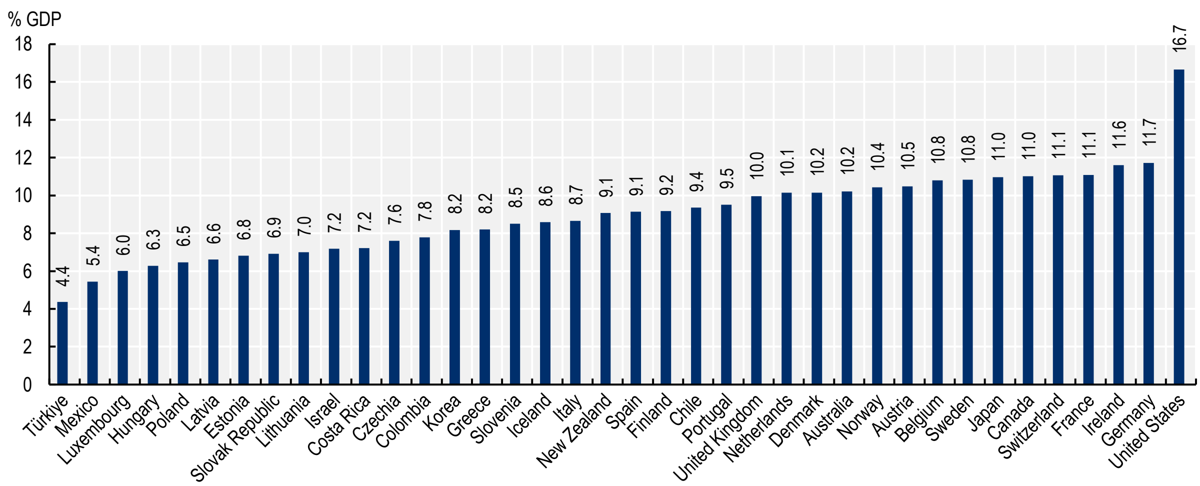
Note: n.a. not available as countries did not provide responses to the question. The qualification of health system features is based on responses to the OECD Health Systems Characteristics survey (see Annex A).

Annex 2.A. Clustering and sensitivity analyses

Input and output variables

The cross-country comparison of health spending as a share of GDP – the input variable of this set of analysis – is shown in Annex Figure 2.A.1.

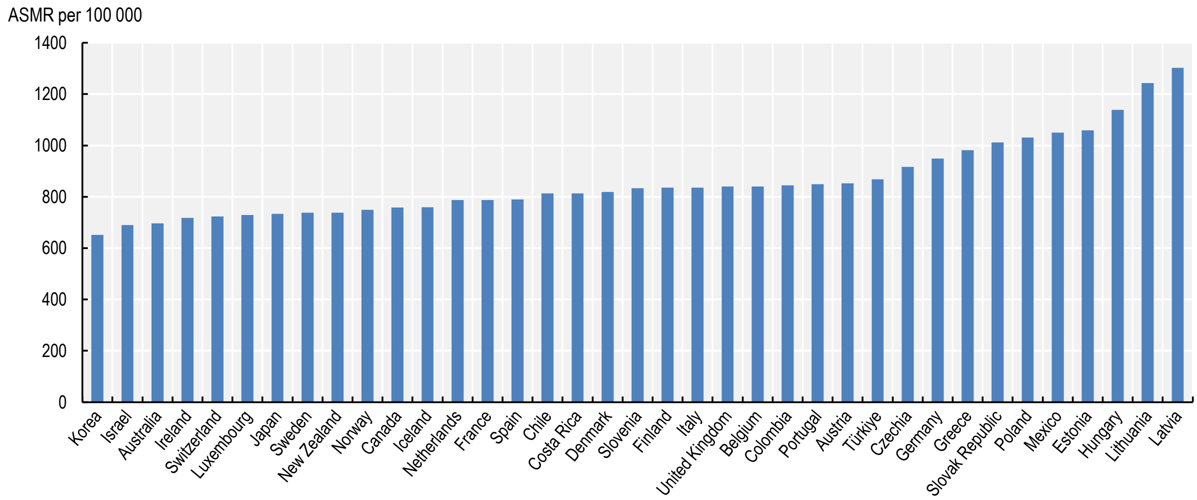
Annex Figure 2.A.1. Health spending as a share of GDP by country, 2019



Note: instead of GDP, GNI is used for Luxembourg and GNI * for Ireland.
Source: OECD Health Statistics 2024.

The cross-country comparison of age standardised mortality rates per 100 000 population – the output variable of this set of analysis – is shown in Annex Figure 2.A.2.

Annex Figure 2.A.2. Age standardised mortality rates by country, 2019

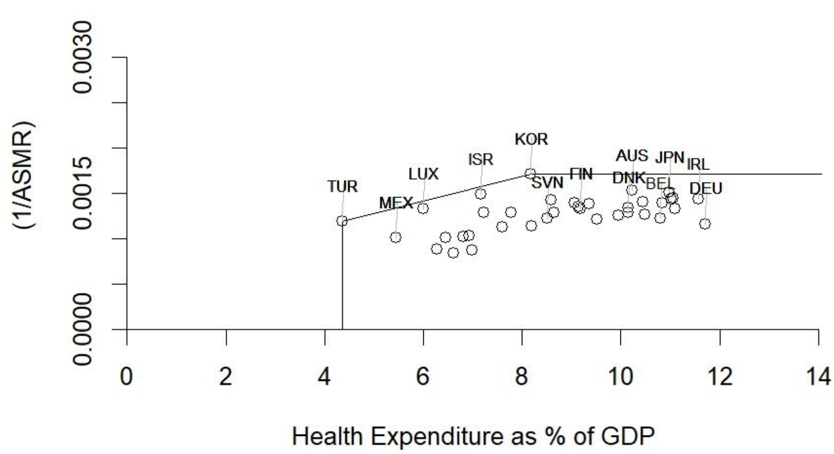


Source: OECD Health Statistics, January 2024.

Efficiency frontier

Annex Figure 2.A.3 shows the efficiency frontier estimated using DEA for 2019 – basic model of output-oriented analysis, one input – one output. It should be noted that the reliability of an efficiency score also depends on the density of observations in the region of the frontier where a country is located. Countries with atypical levels of inputs and outputs tend to be considered as efficient but this result may be just the consequence of the lack of comparable observations (Simar and Wilson, 2007[15]). In this analysis, the "flat" part of the efficiency frontier is computed based on the output for Korea. Countries with a similar or higher level (to that of Korea) of health spending as a share of GDP are compared against the efficiency score for Korea. Countries with a lower level (to that of Korea) of health spending as a share of GDP are compared to the hypothetical frontier represented by the "downwards slope" part of the frontier.

Annex Figure 2.A.3. Efficiency frontier estimated using DEA, 2019



Clustering

Ward's method and Gower's distance

Ward's hierarchical clustering was used to group countries into homogeneous groups. Ward's hierarchical agglomerative clustering is a bottom-up approach where each country starts as its own cluster, and clusters are iteratively merged until a single cluster is formed. The method minimises the total within-cluster variance at each step of the merging process, ensuring that clusters remain as homogenous as possible. Specifically, Ward's method minimises the increase in the total within-cluster sum of squared deviations (the error sum of squares, ESS) when two clusters are merged (Equation 1).

$$ESS(C) = \sum_{i \in C} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{C_j})^2$$

Equation 1

The EES in the context of Ward's method is a measure of the variance within a cluster. It represents the sum of squared deviations of each point in the cluster from the cluster centroid (mean of the points in the cluster). Ward's method merges the two clusters that result in the smallest increase in the total ESS at each step. In the formula x_{ij} is the value of variable *j* for point *i;* $\bar{x}_{C_j}$ is the centroid (mean) of variable j in cluster *C*.

Given that the dataset contains both categorical and continuous variables, Gower's distance was chosen as the distance metric. Gower's distance is a similarity measure that accommodates mixed data types. It standardises continuous variables and treats categorical variables as binary, computing dissimilarities for each variable and then aggregating these into a single distance metric. This makes Gower's distance an appropriate choice when the dataset contains a mix of continuous, ordinal, and nominal variables, ensuring that no data type disproportionately influences the clustering process.

The dendrogram produced by the hierarchical clustering algorithm was used to explore the clustering structure, helping to visualise the merging process. Additionally, the Silhouette score, a metric that evaluates how similar an object is to its own cluster compared to other clusters, was used to assess the quality of the clusters and determine the optimal number of clusters. The silhouette score ranges from -1 to 1, with higher values indicating better-defined clusters.

Quality checks

To ensure the robustness of the clustering results, a series of quality checks were conducted using both cross-validation techniques and feature importance analysis. These analyses ensured that the clusters were both statistically stable and meaningfully interpretable, ensuring that the results were not overly sensitive to small perturbations in the data and that the variables used for clustering were meaningful.

Clustering stability was evaluated using a leave-one-out cross-validation technique. This method involves iteratively excluding one data point (country-year) at a time from the dataset, reapplying the clustering algorithm to the remaining data, and then analysing whether the clusters formed remain consistent. The purpose of this approach is to detect outliers or data points that disproportionately influence the clustering structure. If the exclusion of a single data point leads to major shifts in cluster assignments, this would indicate potential instability in the clustering solution. Conversely, no changes to the clusters would suggest that the solution is robust and not overly dependent on any particular observation.

Additionally, a Random Forest algorithm was employed to assess the importance of variables in determining the cluster structure. Random Forest is a non-parametric ensemble learning method that generates multiple decision trees based on random subsets of the data and then aggregates their predictions. In this context, it was used to evaluate the contribution of each variable to the clustering result.

The variable importance was quantified using two metrics:

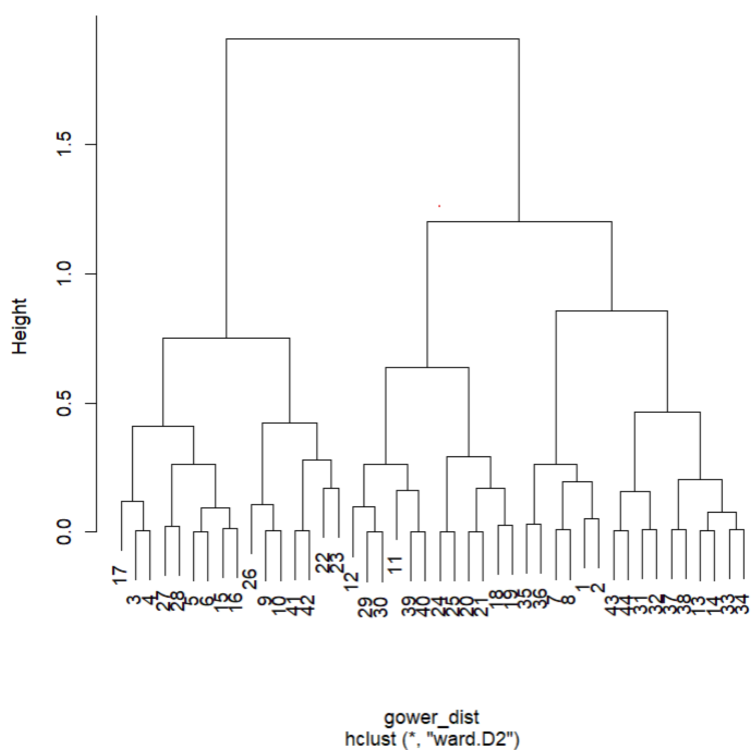
- Mean Decrease Accuracy: This metric measures how much accuracy decreases in the random forest model when a particular variable is randomly permuted, keeping other variables unchanged. A higher mean decrease accuracy indicates that the variable is critical for correct cluster identification.
- Mean Decrease Gini: This metric measures the decrease in node impurity (as quantified by the Gini index) when a variable is used to split the data in a tree. Higher values suggest that the variable plays a significant role in distinguishing between data points in different clusters.

Both metrics were used to rank the importance of variables in defining the clusters. The Random Forest results provided insights into which variables contributed the most to the clustering structure, helping to explain the differentiation between clusters or the distinction of a particular cluster from the rest.

"Pure" statistical clusters

Based on the five indicators used in previous OECD work – that is degree of user choice of basic coverage; degree of private provision of primary care and outpatient specialist services; patient choice of providers; health insurance as a secondary source of coverage ("over the basic" coverage); role of primary care in the health system (gate-keeping) – a hierarchical clustering algorithm was used to create a dendrogram (Annex Figure 2.A.4).

Annex Figure 2.A.4. Dendogram



Seven clusters were then identified as containing elements that were similar among themselves and dissimilar to elements belonging to other groups (Annex Figure 2.A.5).

Annex Figure 2.A.5. Health systems by cluster. Data driven approach

Sensitivity analysis

A series of analyses were performed to understand whether efficiency scores by country – and their relationship with country clusters – were sensitive to the use of:

- two input variables (instead of one input variable base model, that is health expenditure as a share of GDP)
- beta regression model to understand the effect of clusters – instead of visual inspection of efficiency scores of one year
- life expectancy at birth as model output – instead of age-standardised mortality rate

Sensitivity of results to using one additional input variable

Efficiency scores by health system using one additional input variable – GDP per capita; obesity; tobacco consumption; pollution; hospital beds; workforce – were compared against efficiency scores estimated using the base model (Annex Table 2.A.1).

Annex Table 2.A.1. Efficiency scores by country by model using Age Standardised Mortality Rate (ASMR) as output variable

				Model with one additional input variable					
	Year	Base Model (95% Confidence Interval): ASMR – health expenditure as a share of GDP		GDP per capita (in Purchasing Power Parities)	Obesity (% share of population)	Tobacco (% share of population)	Pollution	Hospital beds	Workforce
Australia	2019	**0.84**	(0.77 – 0.88)	0.84	0.84	0.96	0.82	0.98	0.84
Austria	2019	**0.61**	(0.51 – 0.64)	0.61	0.59	0.61	0.78	0.70	0.58
Belgium	2019	**0.56**	(0.46 – 0.6)	0.56	0.55	0.60	0.58	0.69	0.54
Canada	2019	**0.77**	(0.69 – 0.8)	0.77	0.77	0.92	0.92	0.97	0.76
Chile	2019	**0.71**	(0.62 – 0.76)	0.88	0.71	0.72	0.89	0.96	0.92

				Model with one additional input variable					
	Year	Base Model (95% Confidence Interval): ASMR – health expenditure as a share of GDP		GDP per capita (in Purchasing Power Parities)	Obesity (% share of population)	Tobacco (% share of population)	Pollution	Hospital beds	Workforce
Colombia	2019	**0.62**	(0.53 – 0.69)	0.80	0.62	0.64	0.76	0.88	0.90
Costa Rica	2019	**0.69**	(0.61 – 0.76)	0.86	0.70	0.92	0.70	0.87	0.90
Czechia	2019	**0.45**	(0.35 – 0.53)	0.47	0.45	0.46	0.51	0.53	0.54
Denmark	2019	**0.62**	(0.53 – 0.66)	0.62	0.61	0.62	0.75	0.85	0.63
Estonia	2019	**0.42**	(0.3 – 0.5)	0.47	0.41	0.51	0.49	0.50	0.48
Finland	2019	**0.65**	(0.56 – 0.71)	0.65	0.65	0.75	0.86	0.84	0.65
France	2019	**0.68**	(0.59 – 0.71)	0.66	0.65	0.68	0.80	0.80	0.67
Germany	2019	**0.49**	(0.39 – 0.53)	0.49	0.47	0.48	0.64	0.57	0.49
Greece	2019	**0.38**	(0.27 – 0.48)	0.58	0.36	0.40	0.57	0.62	0.69
Hungary	2019	**0.23**	(0.09 – 0.34)	0.29	0.24	0.24	0.29	0.29	0.21
Iceland	2019	**0.72**	(0.63 – 0.79)	0.72	0.72	0.94	0.93	0.93	0.72
Ireland	2019	**0.77**	(0.69 – 0.8)	0.78	0.77	0.83	0.92	0.94	0.76
Israel	2019	**0.88**	(0.8 – 0.94)	0.87	0.87	0.93	0.88	0.92	0.87
Italy	2019	**0.59**	(0.5 – 0.67)	0.63	0.54	0.60	0.66	0.80	0.74
Japan	2019	**0.84**	(0.76 – 0.87)	0.82	0.83	0.83	0.81	0.80	0.84
Korea	2019	**0.92**	(0.85 – 0.99)	0.92	0.85	0.92	0.91	0.92	0.90
Latvia	2019	**0.10**	(0 – 0.2)	0.23	0.10	0.11	0.19	0.20	0.23
Lithuania	2019	**0.12**	(0 – 0.22)	0.18	0.11	0.16	0.22	0.22	0.09
Luxembourg	2019	**0.84**	(0.75 – 0.92)	0.84	0.80	0.93	0.87	0.90	0.84
Mexico	2019	**0.52**	(0.37 – 0.65)	0.81	0.52	0.80	0.57	0.85	0.81
Netherlands	2019	**0.69**	(0.6 – 0.72)	0.69	0.66	0.71	0.76	0.86	0.69
New Zealand	2019	**0.71**	(0.62 – 0.77)	0.73	0.71	0.80	0.69	0.93	0.70
Norway	2019	**0.74**	(0.66 – 0.78)	0.75	0.71	0.92	0.93	0.90	0.75
Poland	2019	**0.43**	(0.31 – 0.53)	0.50	0.41	0.53	0.44	0.51	0.53
Portugal	2019	**0.53**	(0.43 – 0.58)	0.62	0.51	0.59	0.59	0.72	0.54
Slovak Republic	2019	**0.42**	(0.31 – 0.51)	0.49	0.42	0.43	0.49	0.51	0.57
Slovenia	2019	**0.52**	(0.41 – 0.6)	0.60	0.51	0.52	0.59	0.72	0.69
Spain	2019	**0.67**	(0.58 – 0.73)	0.71	0.66	0.68	0.82	0.87	0.81
Sweden	2019	**0.73**	(0.65 – 0.77)	0.73	0.70	0.88	0.92	0.97	0.74
Switzerland	2019	**0.78**	(0.7 – 0.81)	0.79	0.74	0.78	0.94	0.91	0.79

				Model with one additional input variable					
	Year	Base Model (95% Confidence Interval): ASMR – health expenditure as a share of GDP		GDP per capita (in Purchasing Power Parities)	Obesity (% share of population)	Tobacco (% share of population)	Pollution	Hospital beds	Workforce
Türkiye	2019	**0.78**	(0.56 – 0.98)	0.81	0.78	0.81	0.81	0.85	0.81
United Kingdom	2019	**0.59**	(0.5 – 0.63)	0.58	0.59	0.60	0.71	0.83	0.58
OECD average		**0.61**		0.66	0.60	0.67	0.70	0.76	0.67
Pearson correlation with base model				0.93	1.00	0.94	0.94	0.93	0.87

Note: In bold results presented in this report. Instead of GDP, GNI is used for Luxembourg and GNI* for Ireland.

Sensitivity of results to using a beta regression model

To test the robustness of results, efficiency scores for each year from 2016 to 2022 were calculated and then used in a panel data framework to assess whether clusters of health systems help explain observed variations in efficiency.

Initially, a pooled OLS regression was employed to estimate the relationship between various socio-economic, environmental, and health-related factors and efficiency scores. Dummy variables for the eight country clusters were included to assess their effect (*Clusters*). A pooled OLS model was chosen instead of fixed effects due to the invariance of the cluster variables, which would otherwise cause perfect multicollinearity. Testing for time-fixed effects revealed no significant patterns, leading to their exclusion.

The "potential for output increase" form of the efficiency scores was used, which ranges from 1 (most efficient) and above, but usually slightly above 2 (that implies a 100% potential for outcome increase), and applied a logarithmic transformation to reduce skewness and compress the range between 0 and 1. Given these transformed scores, a beta regression model was used, as it is suited to continuous, bounded data and provided a better fit for the efficiency scores. Equation 2 shows the regression formula used.

$$\log(Efficiency_{it}) = \alpha + \beta_0 GDP_{it} + \beta_1\, Education_{it} + \beta_2 Unemployment_{it} + \beta_3 Gini_{it} + \beta_4 Workforce_{it} + \beta_5 Hospital\ Beds_{it} + \beta_6 Obesity_{it} + \beta_7 tobacco_{it} + \beta_8 COVID_{it} + \beta_9 \boldsymbol{Clusters}_i + \varepsilon_{it}$$

Equation 2

Where *i* represents countries an *t* the time in the 2016-22 period*; GDP* is the Gross Domestic Product per capita (in Purchasing Power Parities); *Education* is the percentage of the population aged 25-65 with tertiary education; *Unemployment* is the unemployment rate; *Gini* index is used to measure income inequality; workforce is the rate of healthcare workers per 1 000 population; *Hospital beds* is measured as the rate per 1 000 population; *Obese* is the percentage of the obese population and *tobacco* the percentage of daily smokers aged 15 and above. *Post_COVID-19* is a dichotomous variable that assumes the value of 0 for 2016-18 and 1 for 2019-21. The model presented both heteroskedasticity and autocorrelation, so standard errors were clustered both in time and country, also to account for potential country fixed effects. There was no multicollinearity effect detected (<3 variance inflation factors). There were no outliers or cross-sectional dependence found (studied with Cooks' distance and Pesaran's test, respectively). Annex Table 2.A.2 presents the results of the final model.

Annex Table 2.A.2. Beta regression results

Variable	Estimate	SE	P-values
(Intercept)	2.5 717	0.85687	0**
GDP	-1.89E-05	6.74E-06	0.01**
Education	-3.18E-02	8.59E-03	0***
unemployment	5.02E-03	2.06E-02	0.81
Gini	-5.9267	1.4482	0***
Obese	2.21E-03	3.34E-03	0.51
Tobacco	2.35E-02	5.66E-02	0.68
workforce	6.74E-03	1.28E-02	0.6
Beds	-3.08E-03	1.53E-02	0.84
Post_Covid	3.19E-01	2.03E-01	0.12
Clusters2 (versus cluster 1)	-2.35E-01	2.15E-01	0.27
Clusters3 (versus cluster 1)	-4.26E-02	9.84E-02	0.67
Clusters4 (versus cluster 1)	3.50E-01	3.72E-01	0.35
Clusters5 (versus cluster 1)	1.5779	3.57E-01	0***
Clusters6 (versus cluster 1)	-6.71E-02	3.52E-01	0.85
Clusters7 (versus cluster 1)	1.16E-02	1.56E-01	0.94
Clusters8 (versus cluster 1)	-1.40E-02	1.40E-01	0.92
(phi)	1.99E+01	4.6438	0***
Log-likelihood	*225.5 on 18 Df*		
Pseudo-R2	*0.64*		

Note: Significant result at *0.05, **0.01, ***0.001 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White). The outcome of the model (Efficiency scores in potential for outcome increase) is log transformed.

Sensitivity of results to using a different outcome variable

The sensitivity of results to using a different outcome variable – life expectancy at birth – was tested. Life expectancy has been previously used to proxy health system performance in efficiency analysis and it is widely regarded as one the health system main outputs. Results of the DEA model using life expectancy as the output variable and different specification of inputs can be found in Annex Table 2.A.3.

Annex Table 2.A.3. Efficiency scores by country by model using Life Expectancy (LE) at birth as output variable

				Model with one additional input variable					
	Year	Base Model (95% CI): LE ~ health expenditure as a share of GDP		GDP per capita (in Purchasing Power Parities)	Obesity (% share of population)	Tobacco (% share of population)	Pollution	Hospital beds	Workforce
Australia	2019	0.98	(0.98 – 0.98)	0.98	0.98	0.99	0.98	0.98	0.98
Austria	2019	0.97	(0.96 – 0.97)	0.97	0.97	0.97	0.97	0.97	0.97
Belgium	2019	0.97	(0.96 – 0.97)	0.97	0.97	0.97	0.97	0.97	0.97
Canada	2019	0.97	(0.96 – 0.97)	0.97	0.97	0.98	0.97	0.98	0.97
Chile	2019	0.95	(0.95 – 0.96)	0.98	0.95	0.95	0.95	0.97	0.97
Colombia	2019	0.91	(0.9 – 0.91)	0.98	0.91	0.91	0.91	0.92	0.94
Costa Rica	2019	0.96	(0.95 – 0.97)	0.98	0.96	1.01	0.96	0.98	0.99
Czechia	2019	0.95	(0.94 – 0.95)	0.95	0.94	0.95	0.94	0.95	0.95
Denmark	2019	0.96	(0.96 – 0.97)	0.96	0.96	0.96	0.96	0.97	0.96
Estonia	2019	0.94	(0.93 – 0.95)	0.96	0.94	0.94	0.94	0.94	0.96

				Model with one additional input variable					
	Year	Base Model (95% CI): LE ~ health expenditure as a share of GDP		GDP per capita (in Purchasing Power Parities)	Obesity (% share of population)	Tobacco (% share of population)	Pollution	Hospital beds	Workforce
Finland	2019	0.97	(0.97 – 0.98)	0.97	0.97	0.98	0.98	0.97	0.97
France	2019	0.98	(0.97 – 0.98)	0.98	0.98	0.98	0.98	0.98	0.98
Germany	2019	0.96	(0.95 – 0.96)	0.96	0.96	0.96	0.96	0.96	0.96
Greece	2019	0.97	(0.97 – 0.98)	0.99	0.97	0.97	0.97	0.97	0.99
Hungary	2019	0.91	(0.89 – 0.92)	0.93	0.91	0.91	0.91	0.91	0.93
Iceland	2019	0.99	(0.99 – 0.99)	0.99	0.99	0.99	0.99	0.99	0.99
Ireland	2019	0.98	(0.97 – 0.98)	0.98	0.98	0.98	0.98	0.98	0.98
Israel	2019	0.99	(0.98 – 1)	0.99	0.99	0.99	0.99	0.99	0.99
Italy	2019	0.99	(0.99 – 1)	0.99	0.99	1.00	0.99	0.99	0.99
Japan	2019	1.00	(0.99 – 1)	0.99	0.98	0.99	1.00	0.99	1.00
Korea	2019	0.99	(0.99 – 1)	0.99	0.99	1.00	0.99	0.99	0.99
Latvia	2019	0.90	(0.88 – 0.9)	0.92	0.90	0.90	0.90	0.90	0.92
Lithuania	2019	0.91	(0.9 – 0.91)	0.92	0.91	0.91	0.91	0.91	0.92
Luxembourg	2019	0.99	(0.97 – 1)	0.98	0.98	0.98	0.98	0.98	0.98
Mexico	2019	0.90	(0.88 – 0.92)	0.98	0.90	0.97	0.91	0.97	0.98
Netherlands	2019	0.97	(0.97 – 0.98)	0.97	0.97	0.97	0.97	0.97	0.97
New Zealand	2019	0.97	(0.97 – 0.98)	0.97	0.97	0.98	0.97	0.98	0.97
Norway	2019	0.98	(0.98 – 0.98)	0.98	0.98	0.99	0.98	0.98	0.98
Poland	2019	0.93	(0.92 – 0.94)	0.95	0.93	0.93	0.93	0.93	0.95
Portugal	2019	0.97	(0.97 – 0.97)	0.98	0.97	0.97	0.97	0.97	0.97
Slovak Republic	2019	0.93	(0.92 – 0.93)	0.94	0.93	0.93	0.93	0.93	0.95
Slovenia	2019	0.97	(0.97 – 0.97)	0.97	0.97	0.97	0.97	0.97	0.97
Spain	2019	1.00	(0.99 – 1)	0.99	1.00	1.00	1.00	0.99	0.99
Sweden	2019	0.98	(0.98 – 0.99)	0.98	0.98	0.99	0.98	0.99	0.98
Switzerland	2019	0.99	(0.99 – 0.99)	0.99	0.99	0.99	0.99	1.00	0.99
Türkiye	2019	0.97	(0.92 – 1)	0.98	0.97	0.97	0.97	0.98	0.98
United Kingdom	2019	0.96	(0.96 – 0.97)	0.96	0.96	0.96	0.96	0.97	0.96
OECD average		0.96		0.97	0.96	0.97	0.96	0.97	0.97
Pearson correlation with base model				0.78	1.00	0.88	1.00	0.90	0.87

Note: Instead of GDP, GNI is used for Luxembourg and GNI* for Ireland.

Notes

[1] The term "big bang healthcare reform" refers to large scale changes swiftly implemented (Tuohy, 2011[17]), such as the National Health Service reforms introduced in England in 1991 (Klein, 1995[16]) and the market oriented reforms introduced in New Zealand in 1993 (Ham, 1997[19]).

[2] Data on policies and institutions were available from the questionnaires for the United States. However, reflecting the complexity and variety of the US health system, those data were not used in the sets of analyses discussed in this report. For examples of descriptive analyses that compare the performance of the US healthcare system with that of other high-income countries see reports by the Commonwealth Fund (https://www.commonwealthfund.org/publications/fund-reports/2024/sep/mirror-mirror-2024) and Kaiser

Family Foundation (https://www.kff.org/health-policy-101-international-comparison-of-health-systems/?entry=table-of-contents-introduction).

[3] Using linear programming, Data Envelopment Analysis (DEA) constructs a frontier of the maximum possible output that can be obtained from a given set of inputs (output-oriented efficiency) or of the proportional reduction in input use which is possible keeping output fixed (input-oriented efficiency). This frontier is then used as a benchmark against which the performance of each unit can be assessed. A country's relative distance to the DEA-estimated frontier is interpreted as a measure of potential efficiency gains (Dutu and Sicari, 2016[18]). Compared to parametric approaches to measuring relative efficiency, such as Stochastic Frontier Analysis, DEA does not require assumptions on the underlying production function, even if it still assumes that the latter is common to all units. Moreover, by not making assumptions about the functional form of the relationship between inputs and outputs, DEA is not equipped to provide any conclusions on the expected change in outputs to a marginal change in inputs. In this line, on its own, DEA is not meant to inform how to improve efficiency. Instead, it informs about how far a unit is from the most efficiency possible at that level of input.

[4] For Luxembourg, Gross National Income (GNI) was used to account for compensation of cross-border workers. For Ireland, GNI modified was used to exclude net profits of companies that have been sent abroad, depreciation on Intellectual Property and on leased aircraft, and net income of redomiciled Public Limited Companies.

[5] Age standardised mortality rates in 2019 were used to limit the bias due to the COVID-19 pandemic. In sensitivity analysis, a longer time window (2016-22) was used.

References

Dutu, R. and P. Sicari (2016), "Public Spending Efficiency in the OECD: Benchmarking Health Care, Education and General Administration", *OECD Economics Department Working Papers*, No. 1278, OECD Publishing, Paris, https://doi.org/10.1787/5jm3st732jnq-en. [18]

Gallet, C. and H. Doucouliagos (2017), "The impact of healthcare spending on health outcomes: A meta-regression analysis", *Social Science & Medicine*, Vol. 179, pp. 9-17, https://doi.org/10.1016/j.socscimed.2017.02.024. [11]

Garcia-Escribano, M., T. Mogues and P. Juarros (2022), "Patterns and Drivers of Health Spending Efficiency", *IMF Working Papers*, Vol. 2022/048, p. 1, https://doi.org/10.5089/9798400204388.001. [10]

Ham, C. (1997), "Reforming the New Zealand health reforms", *BMJ*, Vol. 314/7098, pp. 1844-1844, https://doi.org/10.1136/bmj.314.7098.1844. [19]

Hotelling, H. (1933), "Analysis of a complex of statistical variables into principal components.", *Journal of Educational Psychology*, Vol. 24/6, pp. 417-441, https://doi.org/10.1037/h0071325. [2]

Jacobs, R., P. Smith and A. Street (2006), *Measuring efficiency in health care*, Cambridge University Press. [3]

Joumard, I., C. André and C. Nicq (2010), "Health Care Systems: Efficiency and Institutions", *OECD Economics Department Working Papers*, No. 769, OECD Publishing, Paris, https://doi.org/10.1787/5kmfp51f5f9t-en. [1]

Klein, R. (1995), “Big Bang Health Care Reform - Does it work? The Case of Britain’s 1991 National Health Service Reforms”, *The Milbank Quarterly*, Vol. 73/3, pp. 299-337. [16]

Manavgat, G. and M. Audibert (2024), “Healthcare system efficiency and drivers: Re-evaluation of OECD countries for COVID-19”, *SSM - Health Systems*, Vol. 2, p. 100003, https://doi.org/10.1016/j.ssmhs.2023.100003. [8]

Medeiros, J. and C. Schwierz (2015), “Efficiency estimates of health care systems”, European Commission Publications Office, https://data.europa.eu/doi/10.2765/49924. [6]

Moran, V., M. Suhrcke and E. Nolte (2023), “Exploring the association between primary care efficiency and health system characteristics across European countries: a two-stage data envelopment analysis”, *BMC Health Serv Res*, Vol. 23/1, https://doi.org/10.1186/s12913-023-10369-y. [5]

OECD (2020), *Realising the Potential of Primary Health Care*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/a92adee4-en. [14]

OECD (2017), “Life expectancy at birth”, in *Health at a Glance 2017: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/health_glance-2017-6-en. [9]

OECD (2016), *Better Ways to Pay for Health Care*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264258211-en. [13]

OECD/WHO (2014), *Paying for Performance in Health Care: Implications for Health System Performance and Accountability*, Open University Press - McGraw-Hill, Buckingham, https://doi.org/10.1787/9789264224568-en. [12]

Retzlaff-Roberts, D., C. Chang and R. Rubin (2004), “Technical efficiency in the use of health care resources: a comparison of OECD countries”, *Health Policy*, Vol. 69/1, pp. 55-72, https://doi.org/10.1016/j.healthpol.2003.12.002. [7]

Sicari, P. and D. Sutherland (2023), “Health sector performance and efficiency in Ireland”, *OECD Economics Department Working Papers*, No. 1750, OECD Publishing, Paris, https://doi.org/10.1787/6a000bf1-en. [4]

Simar, L. and P. Wilson (2007), “Estimation and inference in two-stage, semi-parametric models of production processes”, *Journal of Econometrics*, Vol. 136/1, pp. 31-64, https://doi.org/10.1016/j.jeconom.2005.07.009. [15]

Tuohy, C. (2011), “American Health Reform in Comparative Perspective: Big Bang, Blueprint, or Mosaic?”, *Journal of Health Politics, Policy and Law*, Vol. 36/3, pp. 571-576, https://doi.org/10.1215/03616878-1271279. [17]

3 Do financial incentives to providers improve performance?

This chapter examines the relationship between healthcare system design – with a particular focus on financial incentives and payment methods – and treatable mortality rates across OECD countries. Using cluster analysis, countries were grouped into three clusters based on their approaches to provider financial incentives and payment systems. The analysis found that health systems using strong financial incentives for quality to providers showed lower treatable mortality rates compared to systems with weak/limited quality incentives. This difference was particularly significant in systems that combined weak quality incentives with fee-for-service payment methods. However, these relationships should not be interpreted as causal due to the lack of an appropriate counterfactual and potential unmeasured system features.

This chapter looks at whether differences in treatable mortality rates across countries can be explained by clustering health systems based on actionable policy levers such as financial incentives to providers to improve healthcare quality and volume incentives embedded in physicians' payment methods.

Avoidable deaths are categorised as those that are either preventable or treatable. A death is considered preventable if it can be avoided through effective public health and primary prevention interventions. On the other hand, a treatable death is a premature death which could be avoided through timely and effective healthcare interventions, including secondary prevention (Nolte and McKee, 2008[1]) (Tobias and Yeh, 2009[2]).

While preventable deaths indicate the state of public health, treatable deaths reflect the availability, accessibility, and quality of healthcare interventions.

In this set of analyses, treatable mortality[1] was used as outcome variable, assuming that a lower rate of deaths amenable to healthcare can indicate an improvement in quality of care. Treatable mortality was identified by OECD and WHO as a key performance indicator that works as tracer for access to and quality of care (Figueras et al., 2023[3]).

Financial incentives to providers have been associated with measures of better access and higher quality of healthcare. An early review by the Cochrane Collaboration based on 4 systematic reviews of 32 studies concluded that financial incentives are effective at changing healthcare professional practice, but, given the paucity of studies, it is still unclear whether they contribute to improved patient outcomes (Flodgren et al., 2011[4]). Along similar lines, a more recent systematic review of studies conducted by (Heider and Mang, 2020[5]) indicated the positive effects on the quality of care of certain types of financial incentives provided to healthcare staff, although it is noteworthy that other studies are less clear with respect to the direction of the effect (for the United Kingdom see (Mandavia et al., 2017[6])). With respect to the use of mixed payment schemes, a growing body of literature has previously demonstrated that combining capitation with fee-for-service (FFS) incentives mitigates significantly the under provision of medical services observed with capitation and the overprovision of services observed with FFS (Brosig-Koch et al., 2013[7]). Bundled payment schemes consistently report increases in efficiency and corresponding cost savings of health systems (Feldhaus and Mathauer, 2018[8]). Regarding patient outcomes, studies have shown that blended capitation payment is associated with improvements in some aspects of diabetes care (Bamimore et al., 2021[9]). Showing similar results but taking the debate a step further, Li et al. conclude that when patients are characterised by lower disease severity and resource consumption is relatively small, FFS may be a more suitable payment method (Li et al., 2022[10]).

Following these findings, we visually analysed the relationship between indicators constructed using country responses to the HSC survey (see Annex A) and the output variable (treatable mortality). This analysis helped identify five indicators to which the output variable is particularly sensitive: volume incentives embedded in physicians' payment schemes, volume incentives embedded in hospital payment schemes, financial incentives for healthcare quality, recruitment and remuneration of hospital staff and the degree of regulation of prices/fees for providers. However, recruitment and remuneration of hospital staff was not used to group health systems due to low face validity (Table 3.1).

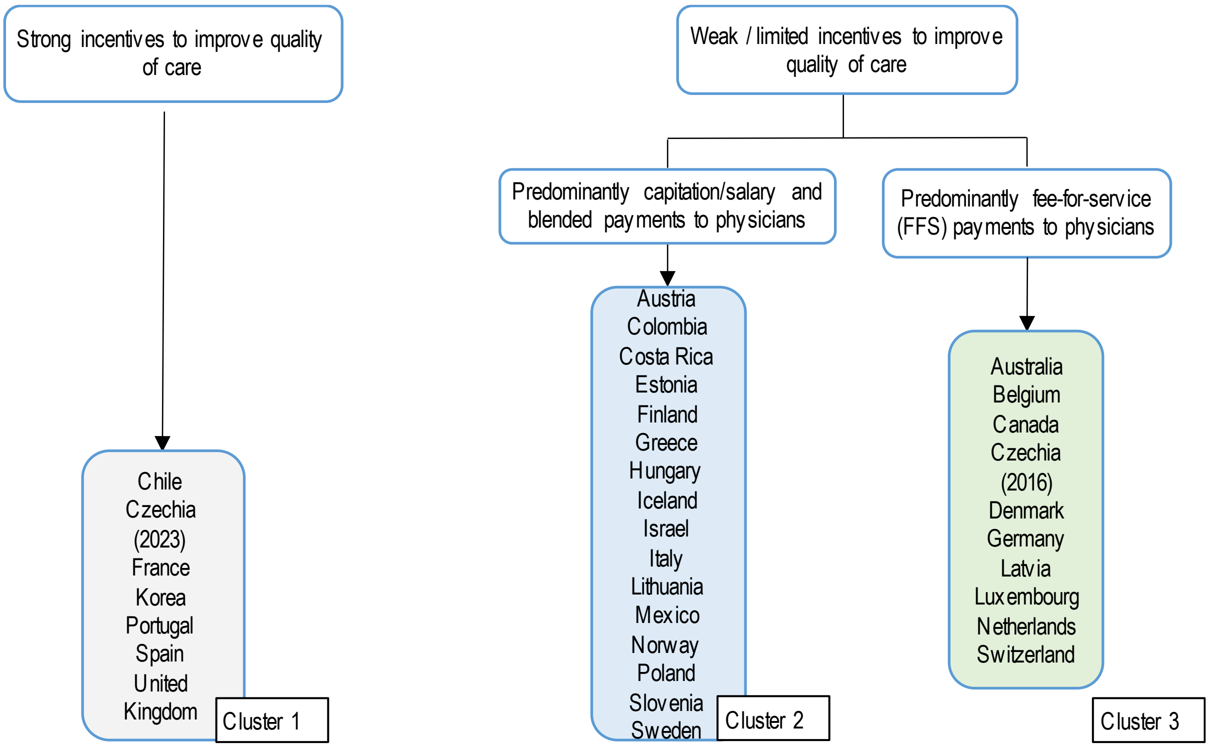
Table 3.1. Description of the indicators used for a cluster analysis on access to high-quality services

Indicator	Description
Volume incentives embedded in physician payment schemes	Evaluates the predominant mode of payment for physicians in primary care, community, and hospitals (i.e. fee-for-service, capitation, salary), as a proxy for incentives to generate volumes of services. A higher value of the indicators means a larger use of fee-for-service.
Volume incentives embedded in hospital payment schemes	Indicates the likely impact of hospitals' predominant payment method (i.e. line-item budget, per diem, global budget, DRGs, per procedure) on volume of care. A higher value of the indicator means a larger use of activity-based payments
Financial incentives for healthcare quality	Shows whether financial bonuses for primary care physicians, specialists and hospitals are provided, related to preventive care, management of chronic diseases, patient satisfaction, uptake of IT services. A higher value of the indicator means a larger use of pay-for-performance
Degree of price/fee regulation	Indicates the degree of price regulation of payments to providers by purchasers. A higher value means that prices are set unilaterally by purchasers at central level

A data driven cluster analysis based on these indicators (see Annex 3.A) indicated that OECD health systems[2] can be grouped into three clusters mainly based on financial incentives to providers to improve quality of care and volume incentives embedded in physicians' payment methods. Expert judgement confirmed the face validity of groups that display the following key features (Figure 3.1):

- Cluster 1: A first group of countries – Chile, Czechia (2023), France, Korea, Portugal, Spain and the United Kingdom – use large financial incentives to providers to improve quality of care.
- Cluster 2: A second group of countries – Austria, Colombia, Costa Rica, Estonia, Finland, Greece, Hungary, Iceland, Israel, Italy, Lithuania, Mexico, Norway, Poland, Slovenia and Sweden are characterised by weak/limited incentives for quality and use blended payment arrangements for physicians (i.e. FFS and capitation).
- Cluster 3: A third group consisting of Australia, Belgium, Canada, Czechia (2016), Denmark, Germany, Latvia, Luxembourg, the Netherlands and Switzerland use weak/limited incentives for quality; however, in contrast to the previous group, their health system is characterised by the use of fee-for service as the predominant payment method for physicians.

Figure 3.1. Groups of health systems with similar financial incentives to providers to improve quality of care and physicians' payment methods



Statistical methods

A pooled panel regression model was used to understand whether clusters of health systems sharing similar financial incentives and payment methods to providers help explain differences in treatable mortality – the output variable.

The model controlled for lifestyle, environmental and socio-economic factors (Table 3.2). These variables are consistent with those included in previous empirical analyses and reflect the key determinants of health outcomes as identified in the relevant literature. Furthermore, it was assumed that if characteristics for a given country did not change between 2016 and 2023, then the health system was assigned to the same cluster for the whole period in study. The model also controlled for two features that were identified as non-actionable from a policy perspective in the absence of large-scale structural reform: the overall type of coverage (residence-based/single payer versus multiple insurers) (Paris et al., 2016[11]) and the degree of decentralisation of spending autonomy in health (Dougherty and Phillips, 2019[12]). Finally, a dummy variable was used to capture the impact of COVID-19.

Table 3.2. Control variables used in the panel regression model

Variable	Indicator	Reference
Wealth	GDP per capita (PPP constant)*	(Hajat et al., 2010[13])
Economic activity	Unemployment rate	(Clemens, Popham and Boyle, 2014[14])
Inequality	Gini index of household income distribution	(Kondo et al., 2009[15])
Education	% of the population 25-65 with tertiary education	(Balaj et al., 2024[16])
Risk factors	• Percentage of obese population • Percentage of population 15+ daily smokers	(Wang et al., 2014[17]) (Glei, Lee and Weinstein, 2022[18])
Environmental hazard	• Pollution (PM2.5 kg per km2)	(Fuller et al., 2022[19])

Note: GNI – instead of GDP – was used for Luxembourg.

Key findings

In line with expectations, the panel regression analysis indicated that lower treatable mortality rates are reported for health systems with higher GDP per capita, higher level of educational attainment and lower obesity rates. Treatable mortality rates were lower in the “multiple insurers and high decentralisation” group of health systems as compared to the “residence-based/single payer and high decentralisation” group of health systems (Table 3.3).

Compared to cluster 1, where health systems use strong incentives for quality, health systems in cluster 2 and cluster 3, which are characterised by weak/limited incentives for quality, reported higher treatable mortality rates. However, the finding for cluster 2 was not statistically significant (Table 3.3).

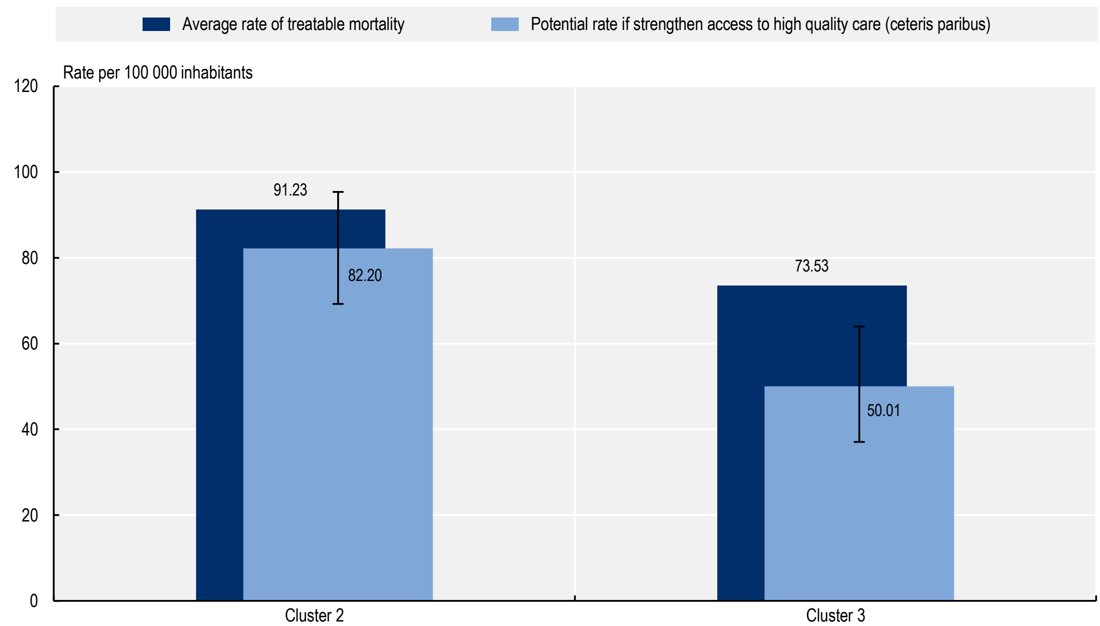
Table 3.3. Variables with a statistically significant coefficient in the regression model

Variable	Estimate	Standard Error	P-value
GDP per capita	-9.25 E-06	2.16 E-06	0.000***
Education	-0.01	0.002	0.000***
Obesity	0.029	0.007	0.001**
NonModChar3 (vs. NonModChar1)	-0.18	0.088	0.06’
Cluster 2 (versus cluster 1)	0.092	0.07	0.19
Cluster 3 (versus cluster 1)	0.33	0.09	0.000***

Note: Significant result at ‘0.1, *0.01, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered at country level. The outcome of the model is log transformed. The model is also controlled for level of unemployment, Gini, PM2.5, tobacco, COVID-19 and non-modifiable characteristics. The full model, together with the functional form, estimation method and assumption testing can be found in Annex 3.A.

Health systems in cluster 3, which combine weak/limited incentives for quality with a predominant use of fee-for-service to pay physicians, reported a statistically significant potential for outcome increase (Figure 3.2).

Figure 3.2. Average difference in treatable mortality rates by cluster



Note: 95% confidence intervals are displayed.

Sensitivity analyses

Sensitivity analyses were conducted to assess the robustness of results to adding covariates such as hospital beds per 1 000 population and workforce per 1 000 population, time fixed effect and year as a numeric variable (see Annex 3.A). The direction and significance of coefficients remained robust across these specifications. However, hospital beds and workforce – metrics that capture capacity – influence the effect of the non-modifiable and modifiable characteristics (clusters) variables.

Financial incentives to providers to increase quality of care are associated with better performance

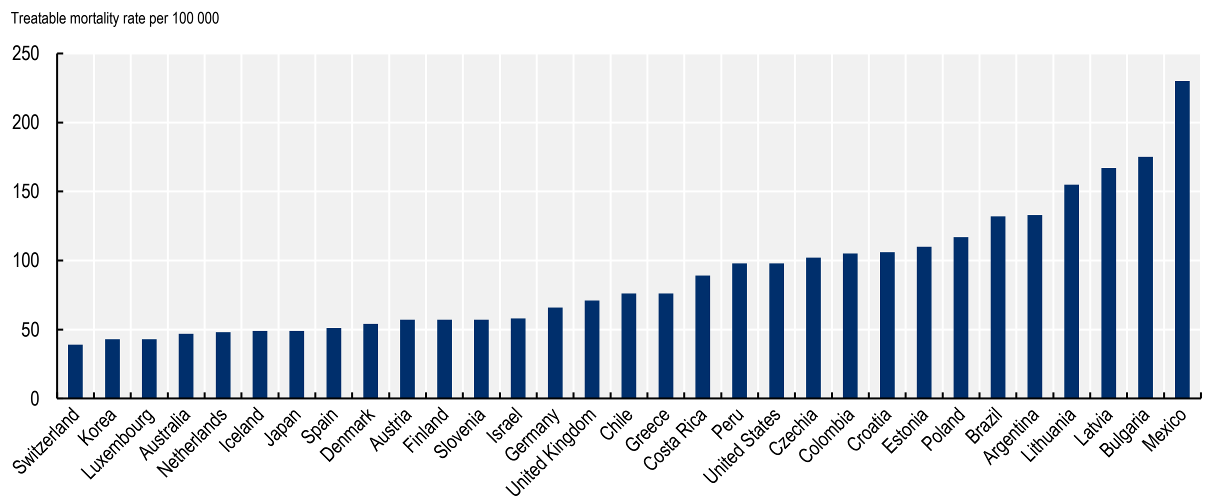
Because of the lack of an appropriate counterfactual, the relationship between the key variables in this analysis should not be interpreted as causal. Moreover, it is possible that the observed effect of clusters on the outcome variable may reflect underlying health system features that were not captured or controlled for in the analysis. Despite these limitations, the results suggest an important association: higher use of financial incentives for providers to enhance quality is linked to a reduction in deaths that should be preventable with high-quality, timely and effective healthcare.

Annex 3.A. Clustering and sensitivity analyses

Output variable

The cross-country comparison of treatable mortality rates per 100 000 population – the output variable of this set of analysis – is shown in Annex Figure 3.A.1.

Annex Figure 3.A.1. Treatable mortality rates by country, 2021



Source: OECD Health Statistics, January 2024.

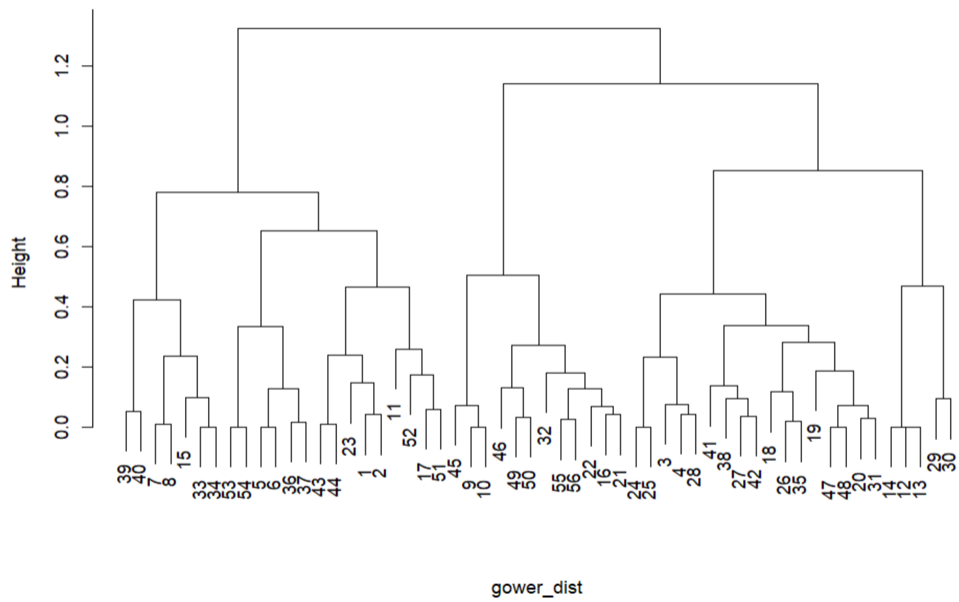
Clustering

Ward's method was employed to group countries into clusters based on four indicators that could be influenced by actionable policy levers: volume incentives embedded in physician payment schemes; volume incentives embedded in hospital payment schemes; financial incentives to increase healthcare quality; intensity of price/fee regulation (Annex Table 3.A.1). A hierarchical clustering algorithm was used to create a dendrogram (Annex Figure 3.A.2).

Annex Table 3.A.1. Score of indicators by cluster

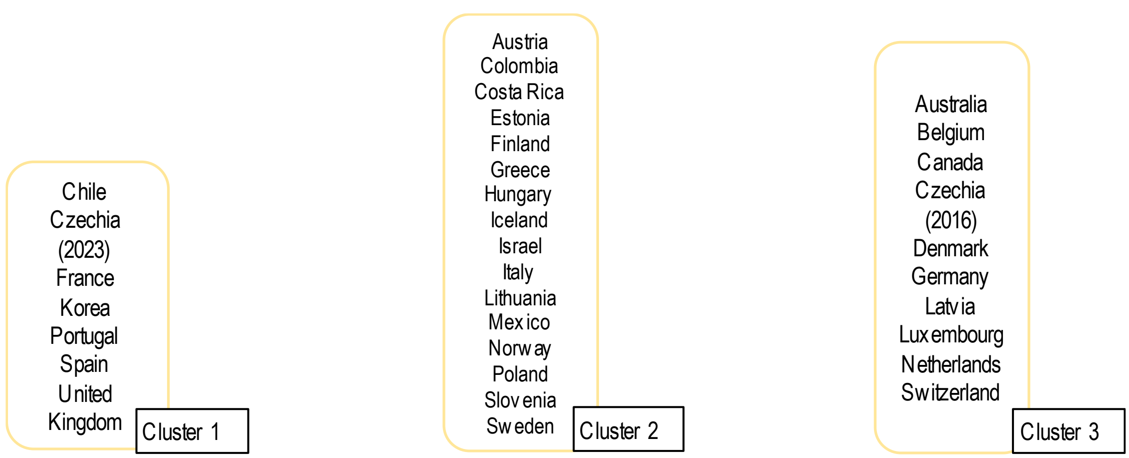
Indicator	Mean		
	Cluster 1	Cluster 2	Cluster 3
Volume incentives embedded in physician payment schemes	0.30	0.19	0.65
Volume incentives embedded in hospital payment schemes	0.65	0.54	0.60
Financial incentives to increase healthcare quality	0.90	0.07	0.26
Intensity of price/fee regulation	0.79	0.76	0.72

Annex Figure 3.A.2. Dendogram



Three clusters were then identified as containing elements that were similar among themselves and dissimilar to elements belonging to other groups (Annex Figure 3.A.3).

Annex Figure 3.A.3. Health systems by cluster. Data driven approach



Statistical approach

Pooled ordinary least squares regression

A pooled ordinary least squares regression was used to estimate the relationship between various socio-economic, environmental, and health-related factors and treatable mortality, defined as the mortality amenable to quality healthcare. The dependent variables, treatable mortality rate is operationalised as the logarithm to improve model fit. The explanatory variables include Gross Domestic Product (*GDP*) per capita, adjusted for purchasing power parity, which serves as a control for wealth of the country. The unemployment rate was also included, in order to capture the economic activity. The Gini index was used to measure income inequality. Educational attainment was represented by the percentage of the population aged 25-65 with tertiary education, which reflects the overall education level of the country. Health-related behaviours were captured through variables such as the percentage of the obese population and the percentage of daily smokers aged 15 and above. Environmental exposure was represented by the mean annual concentration of particulate matter (*PM2.5 in mg per m^3)* per squared kilometre.

In addition to the model variables outlined above, the model included care capacity indicators of hospital beds per 1 000 population (*Beds*) and Primary care practitioners (General practitioners, paediatricians and gynaecologist) per 1 000 population (*PCP*). In time, capacity indicators (rate of workforce and hospital beds per 1 000 population) were included only as sensitive analysis.

Additionally, Nonmodifiable health system characteristics ($NonModChar$) were included to account for aspects of the health system that are not subject to change in the short term (Annex Table 3.A.2).

Annex Table 3.A.2. Non-modifiable health system characteristics by country

	Spending autonomy in health	
Coverage	High decentralisation	Low decentralisation
Residence-based – single payer	Australia, Canada, Denmark, Finland, Italy, New Zealand, Spain, Sweden Türkiye, United Kingdom	Costa Rica, Estonia, Greece, Hungary, Iceland, Ireland, Korea, Latvia, Lithuania, Luxembourg Norway, Poland, Portugal, Slovenia
Multiple insurers	Chile, Colombia, France, Germany, Israel, Japan, Netherlands, Slovak Republic, the United States	Austria, Belgium, Czechia, Mexico, Switzerland

Source: OECD Health Systems Characteristics Survey; Dougherty, S. and L. Phillips (2019[12]), "The spending power of sub-national decision makers across five policy sectors", https://doi.org/10.1787/8955021f-en.

Finally, the model used dummy variables ($Clusters$) to represent the three clusters of countries (Equation 3). The analysis is conducted using the "plm" package in R.

$$\log(TreatMort_{it}) =$$
$$\alpha + \beta_0 GDP_{it} + \beta_1\, Education_{it} + \beta_2 Unemployment_{it} + \beta_3 Gini_{it} + \beta_4 Pollution + \beta_5 Obesity_{it} + \beta_6 tobacco_{it}$$
$$+\beta_7 COVID_{it} + \beta_8 NonModChar + \beta_9 \boldsymbol{Clusters} + \varepsilon_{it}$$

Equation 3

With i representing OECD countries and t the year in the 2016-22 period.

Assumption testing

Several tests were conducted to assess the underlying assumptions of our panel models. We first performed a Chow test for poolability using the pooltest function (R package plm), which examines whether the panel data model benefits from individual effects beyond the pooled model. For both models, this test suggested potential individual effects across entities, significant at 5% error. Due to the static nature of our interest variables, we could not be adequately controlled by a fixed effects model at country level without losing these variables from the analysis, thus becoming one of the limitations of our approach. Nevertheless, we clustered our standard errors at the country level to account for potential country level fixed effects and not overestimate the significance of the results. Similarly, we tested the significance of and included time fixed effects when appropriate by calculating F Test for Individual and/or Time Effects. In the final models, time fixed effects were not significant when adding the COVID-19 variable.

Additionally, Pesaran's test of cross-sectional dependence was applied to investigate the independence of errors across panel units. For heteroskedasticity, a Breusch-Pagan test indicated the presence of heteroskedastic error terms in both models, which was addressed by computing robust standard errors using the Arellano method. The Breusch-Godfrey test was conducted for autocorrelation, which checks for serial correlation in the residuals, and assessed multicollinearity among independent variables using the Variance Inflation Factor. The results of the main model are summarised in detail in Annex Table 3.A.3. The variables in the model were removed one by one, with the overall model conclusions remaining strong to the different specifications. The final model corresponds to the one showing the highest R squared.

Annex Table 3.A.3. Panel regression results

Variable	Estimate	SE	P-values
(Intercept)	3.88E+00	2.71E-01	2.14E-14***
GDP	-9.25E-06	2.18E-06	2.21E-04***
Education	-1.31E-02	3.01E-03	1.58E-04***
Unemployment	-1.81E-02	8.58E-03	4.34E-02*
Gini	1.29E+00	6.45E-01	5.48E-02
PM2.5_km	-1.52E-04	1.91E-04	4.32E-01
Obese	2.92E-02	7.85E-03	8.86E-04***
Tobacco	1.61E-02	8.34E-03	6.40E-02
Post_Covid years	4.22E-02	4.02E-02	3.02E-01
NonModChar2 (vs. NonModChar1)	9.70E-02	9.59E-02	3.20E-01
NonModChar3 (vs. NonModChar1)	-1.76E-01	8.83E-02	5.59E-02’
NonModChar4 (vs. NonModChar1)	-1.06E-01	9.64E-02	2.83E-01
Cluster 2 (versus cluster 1)	9.92E-02	7.36E-02	1.89E-01
Cluster 3 (versus cluster 1)	3.26E-01	9.98E-02	2.88E-03**
R2	*0.89*		
R2adj	*0.84*		
F-statistic	*17.7085 on 13 and 28 DF, p-value 3.72E-10*		

Note: Significant result at ‘0.1, *0.05, **0.01, ***0.001 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.
NonModChar1: Residence-based or single payer + Relatively high spending power in health of subnational decision makers.
NonModChar2: Residence-based or single payer + Relatively low spending power in health of subnational decision makers.
NonModChar3: Multiple insurers + Relatively high spending power in health of subnational decision makers.
NonModChar4: Multiple insurers + Relatively low spending power in health of subnational decision makers.

Sensitivity analyses

The model was tested with various specifications, incorporating variables such as hospital beds per 1 000 population and workforce per 1 000 population, time fixed effects and year as a numeric variable.

The direction and significance of the coefficients remained robust across these specifications. However, healthcare capacity variables (workforce and beds) influence the effect of the nonmodifiable and modifiable characteristics (clusters) variables (Annex Table 3.A.4). These findings are consistent with the expected mediation effect of health system inputs on the relationship between characteristics and output.

Annex Table 3.A.4. Sensitivity analysis: Panel regression results

Variable	Estimate	SE	P-values
(Intercept)	2.86E+00	3.34E-01	4.63E-09***
GDP	-1.46E-05	2.72E-06	1.23E-05***
Education	-1.41E-02	2.30E-03	1.79E-06***
Unemployment	-7.78E-03	6.15E-03	2.17E-01
Workforce	4.66E+00	8.35E-01	7.38E-06***
Beds	-3.00E-04	1.38E-04	3.91E-02*
Gini	2.63E-02	7.15E-03	1.06E-03**
PM2.5_km	-7.02E-03	9.80E-03	4.80E-01
Obese	2.86E+00	3.34E-01	4.63E-09***
Tobacco	-1.46E-05	2.72E-06	1.23E-05***
Workforce	1.45E-01	2.94E-02	7.75E-02
Beds	2.88E-02	3.79E-02	4.23E-05***
Post_Covid years	7.93E-02	7.49E-02	4.54E-01
NonModChar2 (vs. NonModChar1)	-1.45E-01	7.24E-02	2.99E-01
NonModChar3 (vs. NonModChar1)	-2.66E-01	7.70E-02	5.58E-02
NonModChar4 (vs. NonModChar1)	1.00E-01	7.25E-02	1.88E-03**
Cluster 2 (versus cluster 1)	3.67E-01	9.50E-02	1.79E-01
Cluster 3 (versus cluster 1)	1.45E-01	2.94E-02	6.68E-04***
R2	*0.93*		
R2adj	*0.90*		
F-statistic	*25.3665 on 15 and 26 DF, p-value 9.35E-12*		

Note: Significant result at '0.1, *0.05, **0.01, ***0.001 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.
NonModChar1: Residence-based or single payer + Relatively high spending power in health of subnational decision makers.
NonModChar2: Residence-based or single payer + Relatively low spending power in health of subnational decision makers.
NonModChar3: Multiple insurers + Relatively high spending power in health of subnational decision makers.
NonModChar4: Multiple insurers + Relatively low spending power in health of subnational decision makers.

Notes

[1] COVID-19 deaths are not included in the list of causes of treatable mortality.

[2] Ireland, Japan, New Zealand, the Slovak Republic and Türkiye are not included in this analysis as there are missing values for the indicators used to cluster health systems.

References

Balaj, M. et al. (2024), "Effects of education on adult mortality: a global systematic review and meta-analysis", *The Lancet Public Health*, Vol. 9/3, pp. e155-e165, https://doi.org/10.1016/s2468-2667(23)00306-7. [16]

Bamimore, M. et al. (2021), "Quality of Diabetes Care in Blended Fee-for-Service and Blended Capitation Payment Systems", *Canadian Journal of Diabetes*, Vol. 45/3, pp. 261-268.e11, https://doi.org/10.1016/j.jcjd.2020.09.002. [9]

Brosig-Koch, J. et al. (2013), “How to Improve Patient Care? - An Analysis of Capitation, Fee-for-Service, and Mixed Payment Schemes for Physicians”, *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.2278841. [7]

Clemens, T., F. Popham and P. Boyle (2014), “What is the effect of unemployment on all-cause mortality? A cohort study using propensity score matching”, *European Journal of Public Health*, Vol. 25/1, pp. 115-121, https://doi.org/10.1093/eurpub/cku136. [14]

Dougherty, S. and L. Phillips (2019), “The spending power of sub-national decision makers across five policy sectors”, *OECD Working Papers on Fiscal Federalism*, No. 25, OECD Publishing, Paris, https://doi.org/10.1787/8955021f-en. [12]

Feldhaus, I. and I. Mathauer (2018), “Effects of mixed provider payment systems and aligned cost sharing practices on expenditure growth management, efficiency, and equity: a structured review of the literature”, *BMC Health Services Research*, Vol. 18/1, https://doi.org/10.1186/s12913-018-3779-1. [8]

Figueras, J. et al. (eds.) (2023), *Assessing Health System Performance: Proof of Concept for a HSPA Dashboard of Key Indicators*, OECD Publishing, Paris/World Health Organization, Geneva, https://doi.org/10.1787/4e6b28c0-en. [3]

Flodgren, G. et al. (2011), “An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes”, *Cochrane Database of Systematic Reviews*, https://doi.org/10.1002/14651858.cd009255. [4]

Fuller, R. et al. (2022), “Pollution and health: a progress update”, *The Lancet Planetary Health*, Vol. 6/6, pp. e535-e547, https://doi.org/10.1016/s2542-5196(22)00090-0. [19]

Glei, D., C. Lee and M. Weinstein (2022), “Assessment of Mortality Disparities by Wealth Relative to Other Measures of Socioeconomic Status Among US Adults”, *JAMA Network Open*, Vol. 5/4, p. e226547, https://doi.org/10.1001/jamanetworkopen.2022.6547. [18]

Hajat, A. et al. (2010), “Long-Term Effects of Wealth on Mortality and Self-rated Health Status”, *American Journal of Epidemiology*, Vol. 173/2, pp. 192-200, https://doi.org/10.1093/aje/kwq348. [13]

Heider, A. and H. Mang (2020), “Effects of Monetary Incentives in Physician Groups: A Systematic Review of Reviews”, *Applied Health Economics and Health Policy*, Vol. 18/5, pp. 655-667, https://doi.org/10.1007/s40258-020-00572-x. [5]

Kondo, N. et al. (2009), “Income inequality, mortality, and self rated health: meta-analysis of multilevel studies”, *BMJ*, Vol. 339/nov10 2, pp. b4471-b4471, https://doi.org/10.1136/bmj.b4471. [15]

Li, X. et al. (2022), “Effects of fee-for-service, diagnosis-related-group, and mixed payment systems on physicians’ medical service behavior: experimental evidence”, *BMC Health Services Research*, Vol. 22/1, https://doi.org/10.1186/s12913-022-08218-5. [10]

Mandavia, R. et al. (2017), “Effectiveness of UK provider financial incentives on quality of care: a systematic review”, *British Journal of General Practice*, Vol. 67/664, pp. e800-e815, https://doi.org/10.3399/bjgp17x693149. [6]

Nolte, E. and M. McKee (2008), “Measuring the health of nations: updating an earlier analysis”, *Health Affairs*, Vol. 27/1, pp. 58-71, https://doi.org/10.1377/hlthaff.27.1.58. [1]

Paris, V. et al. (2016), "Health care coverage in OECD countries in 2012", *OECD Health Working Papers*, No. 88, OECD Publishing, Paris, https://doi.org/10.1787/5jlz3kbf7pzv-en. [11]

Tobias, M. and L. Yeh (2009), "How much does health care contribute to health gain and to health inequality? Trends in amenable mortality in New Zealand 1981–2004", *Aust N Z Public Health*, Vol. 33/1, pp. 70-78, https://doi.org/10.1111/j.1753-6405.2009.00342.x. [2]

Wang, X. et al. (2014), "Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies", *BMJ*, Vol. 349/jul29 3, pp. g4490-g4490, https://doi.org/10.1136/bmj.g4490. [17]

4 Does a strong primary care system improve performance?

This chapter investigates the relationship between primary care system characteristics and avoidable hospital admissions across OECD countries, focusing on admissions for asthma, chronic obstructive pulmonary disease, and congestive heart failure. Using cluster analysis, countries were grouped into five clusters based on three key primary care features: gatekeeping roles, continuity of care, and financial incentives for providers. The analysis revealed that health systems with strong gatekeeping, high continuity of care, and substantial financial incentives for quality demonstrated lower rates of avoidable hospital admissions compared to other systems. While the findings suggest that primary care oriented systems may help reduce acute deterioration in patients with chronic conditions, the large confidence intervals in the results warrant careful interpretation as the relationship may be influenced by unmeasured health system features or social determinants of health.

This chapter looked at whether differences in avoidable hospital admissions for selected conditions – asthma, chronic obstructive pulmonary disease, and congestive heart failure – can be linked to primary care oriented systems.

A high-performing primary care system, which provides accessible and high-quality services, can improve population health (Stange, Miller and Etz, 2023[1]) and reduce acute deterioration in people living with asthma, chronic obstructive pulmonary disease, or congestive heart failure, widely prevalent long-term conditions. The evidence base for effective treatment is well established, and hospital admissions for these conditions are largely avoidable and are therefore used as a marker of quality and access to primary care (OECD/European Commission, 2024[2]).

There is a long-standing debate on the features of high-quality primary care. When defining a "good" primary care system, there is a broad consensus that such system should be characterised by its accessibility, the provision of comprehensive care tailored to patient needs, with strong co-ordination across all healthcare levels and continuity of care, particularly for individuals living with chronic conditions (OECD, 2020[3]). Primary care delivery should rely on robust relationships between the interdisciplinary primary care team, the patient, their family, and the community. This system must be adept at co-ordinating activities across different levels of healthcare, serving as the primary point of contact, especially for patients with complex care needs. Across the board, there is agreement that a well-functioning primary healthcare system improves care co-ordination and health outcomes and reduces wasteful spending, by limiting unnecessary hospitalisations and their associated costs (OECD, 2020[3]).

Considering the above, Ambulatory Care Sensitive Conditions (ACSC) are widely recognised as an indicator of access, quality and performance of the primary healthcare system (Agency for Healthcare Research and Quality, 2002[4]). ACSCs refer to acute or chronic health conditions that lead to avoidable hospitalisations if not effectively managed early in the outpatient primary care setting. Timely and adequate treatment for ACSCs, delivered at a primary care level is shown to reduce the need for hospital admissions, leading to better healthcare outcomes for patients.

Care and management of ACSCs is complex, often involving different care providers in different settings across the healthcare system, rendering the critical role of the primary care system as focal. Multiple studies have shown that continuity of care and good co-ordination between healthcare professionals are strongly associated with lower rates of ACSC-related hospital admissions (Van Loenen et al., 2015[5]; Van Loenen et al., 2016[6]; Lyhne et al., 2022[7]). Research findings also indicate that hospital bed supply is an important factor in determining the number of ACSC-related hospital admissions. The availability of acute hospital beds in particular may influence admission decisions and lead to unnecessary ACSC-related hospital admissions, even when outpatient care could suffice (O'Cathain et al., 2013[8]; Van Loenen et al., 2016[6]).

In this set of analyses, the age-sex standardised rate of potentially avoidable hospital admissions due to asthma, chronic obstructive pulmonary disease, and congestive heart failure per 100 000 population was used as output variable, based on the assumption that a lower rate can indicate higher quality of and better access to primary care (OECD, 2020[3]). The OECD also looked at potentially avoidable hospital admissions for two additional conditions – diabetes and hypertension. However, variations in coding practices for these conditions limit international comparisons. As a result, OECD discontinued the collection of hospitalisation data on hypertension, and further efforts are underway to improve the accuracy and international comparability of data related to hospital admissions for diabetes. Consequently, the analysis in this chapter focuses on chronic obstructive pulmonary disease, asthma and congestive heart failure.

While hospital-based measures provide valuable insights into system-level impacts, they do not fully capture the direct effects of primary care services on patient health and well-being. Future analyses could incorporate patient outcome variables, such as those defined by the OECD PaRIS (Patient-Reported Indicators Surveys) framework, to provide a more comprehensive evaluation.

Incentives for primary healthcare providers have demonstrated moderate effectiveness in reducing avoidable hospitalisations, particularly by enhancing chronic disease management and improving the quality of preventive care. Some studies suggest that these incentives can align provider efforts with desired care outcomes, leading to partial or positive impacts on care quality (Petersen et al., 2006[54]). However, the overall evidence remains mixed, as the effectiveness of such programmes varies significantly depending on their design, the scale of incentives, and the specific populations targeted. Despite this variability, incentive-based approaches appear to encourage a shift toward more proactive and patient-centred care practices, highlighting their potential when thoughtfully implemented.

There is a range of ways to assess the performance of primary care at system level (World Health Organization and the United Nations Children's Fund (UNICEF), 2022[9]; Kringos et al., 2013[10]; Gumas et al., 2024[11]). In this chapter, considering the key domains used by Barbara Starfield's 4Cs framework (Macinko, Starfield and Shi, 2003[12]; Starfield, Shi and Macinko, 2005[13]) to assess the contribution of primary care to health outcomes, we aimed to identify indicators of primary care for which the data available through the OECD survey on Health System Characteristics cover a large part of OECD countries (Table 4.1): gatekeeping; population registered with a primary care provider and/or with regular doctor to go for care; financial incentives to primary care physicians to provide preventive care, to manage chronic disease and population risk factors, to uptake IT services and related to patient satisfaction. Those indicators are mainly related to the service-delivery process (Kringos et al., 2013[14]).

Table 4.1. Description of indicators used for a cluster analysis on the strength of primary care

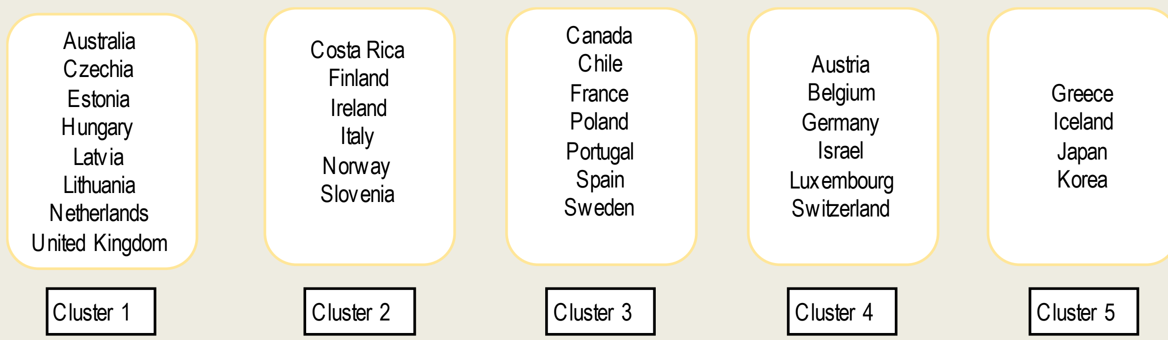
Domain	Indicator	Description
First contact	Role of primary care in the health system (gate-keeping)	Show whether there is a requirement that primary care practitioners serve as gatekeepers to other levels of care. C: strong gatekeeping; B: limited gatekeeping; A: weak/no gatekeeping.
Continuity	Population registered with a primary healthcare provider and/or with a regular doctor to go for care	It captures the share of the population that has a regular doctor to go for care. A: limited part of the population; B: the majority of the population; C: almost the whole population.
Incentives	Financial incentives for primary care physicians	Indicates degree of financial incentives to primary care physicians to provide preventive care, to manage chronic disease and population risk factors, to uptake IT services and related to patient satisfaction. A greater number of positive responses indicate strong financial incentives. A higher score is then assigned.

First, a pure statistical data driven approach to clustering was conducted based on those three indicators (Box 4.1).

Box 4.1. Results of the data driven approach to clustering health systems based on financial incentives for quality of care and payment systems to providers

The data driven approach identified the following five clusters of health systems:

Figure 4.1. Clusters of health systems

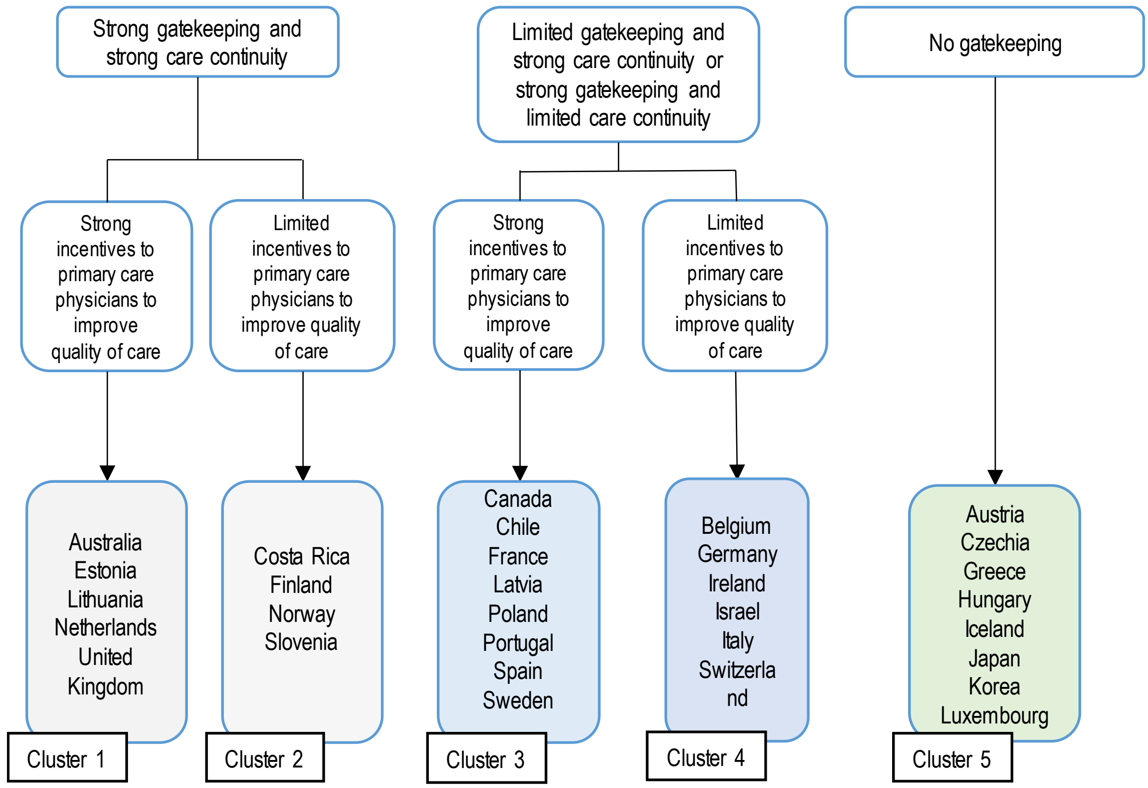


At first sight, some of the clusters look plausible – most of them include countries which are neighbours of each other and/or which have aspects in common. However, there are interlopers that are difficult to explain (e.g. Czechia in cluster 1; Luxembourg in cluster 4).

Based on the data driven approach, some expert judgement was used to ensure that there are meaningful and identifiable policy differences to explain why health systems are grouped together. As a result, five clusters[1] that display the following features were identified (Figure 4.2):

- Cluster 1: Australia, Estonia, Lithuania, the Netherlands and the United Kingdom report a strong gatekeeping role of primary care physicians, a high continuity of care and large financial incentives to primary care physicians.
- Cluster 2: Costa Rica, Finland, Norway and Slovenia report a strong gatekeeping role of primary care physicians and a high continuity of care. In contrast to cluster 1, financial incentives for quality to primary care physicians are generally weaker.
- Cluster 3: Canada, Chile, France, Latvia, Poland, Portugal, Spain and Sweden report either a strong gatekeeping and a limited care continuity or a limited gatekeeping and a strong care continuity as well as strong financial incentives for quality to primary care physicians.
- Cluster 4: Belgium, Germany, Ireland, Israel, Italy and Switzerland report either a strong gatekeeping and limited continuity of care or a limited gatekeeping and a strong continuity of care. In contrast to cluster 3, financial incentives for quality to primary care physicians are generally weaker.
- Cluster 5: Austria, Czechia, Greece, Hungary, Iceland, Japan, Korea and Luxembourg do not report a "gatekeeping" role of primary care physicians.

Figure 4.2. Groups of health systems with similar gate-keeping function, continuity of care and financial incentives for primary care physicians to improve quality of care



Statistical methods

Similarly to the previous set of analyses (see Chapter 3), the model controls for several characteristics that influence health system performance in primary healthcare (Table 4.2). Control variables also accounted for the two features considered non-actionable – that is more complex to change – from a policy perspective: the overall type of coverage (residence-based/single payer versus multiple insurers) (Paris, Devaux and Wei, 2010[15]) and the degree of decentralisation of spending autonomy in health (Dougherty and Phillips, 2019[16]). Finally, a dummy variable was used to capture the impact of COVID-19.

Table 4.2. Control variables used in the panel regression model

Variable	Indicator	Reference
Wealth	GDP per capita (measured in Purchasing Power Parities)	(Macinko, Starfield and Shi, 2003[12])
Inequality	Gini index of household income distribution	
Education	% of the population 25-65 with tertiary education	
Demographics	% of the population 65 years old or older	(Macinko, Starfield and Shi, 2003[12])
Risk factors	• % of obese population • Percentage of population 15+ daily smokers	(Macinko, Starfield and Shi, 2003[12])
Environmental hazard	• Pollution (PM2.5 kg per km2)	
Health system capacity	• Hospitalisation rates (discharges per 100 000 population)/Hospital beds per 1 000 population • Primary care physicians (General practitioners, gynaecologists, paediatricians) per 1 000 population.	(Macinko, Starfield and Shi, 2003[12])
Major disruptions to hospital admissions	• A discrete variable to indicate the years of the COVID-19 pandemic	

Note: GNI – instead of GDP – was used for Luxembourg.

Key findings

In line with expectations, the panel regression analysis indicated that higher avoidable admission rates are reported in countries with higher inequality in income distribution, higher level of educational attainment, and in health systems with a higher number of hospital beds per 1 000 population. Avoidable hospital admissions rates were higher in the “multiple insurers and high decentralisation” group of health systems as compared to the “residence-based/single payer and high decentralisation” group of health systems (Table 4.3).

Cluster 1 – which groups health systems with strong gatekeeping where most of the population regularly sees the same family doctor and large financial incentives for quality to primary care physicians – presents lower avoidable admission rates relative to the other clusters. This difference is statistically significant (Table 4.3).

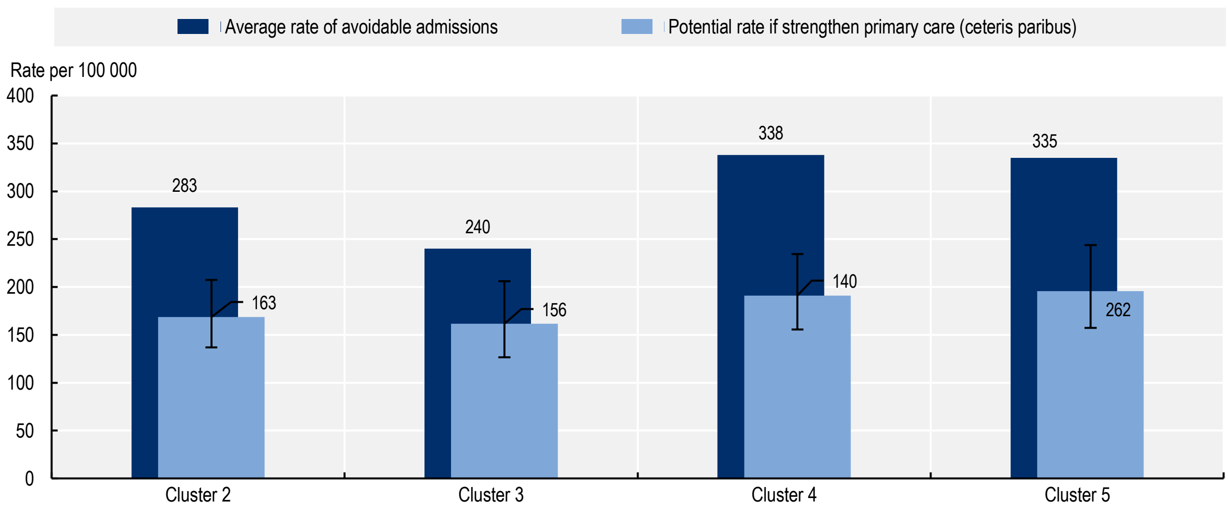
Table 4.3. Variables with a statistical significant coefficient in the regression model

Variable	Estimate	Standard Error	P-value
Education	0.02	0.00	0.00***
Gini	3.29	1.36	0.02*
Obese	0.02	0.01	0.03*
Hospitalization rate	0.00	0.00	0.00***
Primary care physicians	-0.16	0.08	0.05*
Post_Covid	-0.18	0.05	0.00***
NonModChar3 (vs. NonModChar1)	0.72	0.18	0.00***
Cluster 2 (versus cluster 1)	0.52	0.11	0.00***
Cluster 3 (versus cluster 1)	0.40	0.12	0.00**
Cluster 4 (versus cluster 1)	0.57	0.10	0.00***
Cluster 5 (versus cluster 1)	0.54	0.11	0.00***

Note: Significant result at *0.05, **0.01, ***0.001 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered at country level. The outcome of the is log transformed. The model is also controlled for health behaviours, environmental hazards, rate of PCPs, COVID-19 years and non-modifiable characteristics. The full model, together with the functional form, estimation method and assumption testing can be found in Annex 4.A.

Figure 4.3 shows the potential increase in performance for the other clusters of health systems should they adopt primary care oriented policies similar to the one reported for health systems in cluster 1. The large confidence intervals call for caution when interpreting those results.

Figure 4.3. Potential average decrease in avoidable hospital admissions by cluster



Note: 95% confidence intervals are displayed.

Sensitivity analyses

Sensitivity analyses were conducted to assess the robustness of results to using time fixed effects and to different output variables (see Annex 4.A). The direction and significance of coefficients of control variables remained robust across those specifications. However, the link between clusters and the output variable was not confirmed when using congestive heart failure admission rates alone, while only cluster 2 and 4 had a statistically significant effect when using asthma hospital admissions as the output variable.

Care continuity, strong gatekeeping, and large financial incentives for quality of care are related to lower rates of avoidable admissions

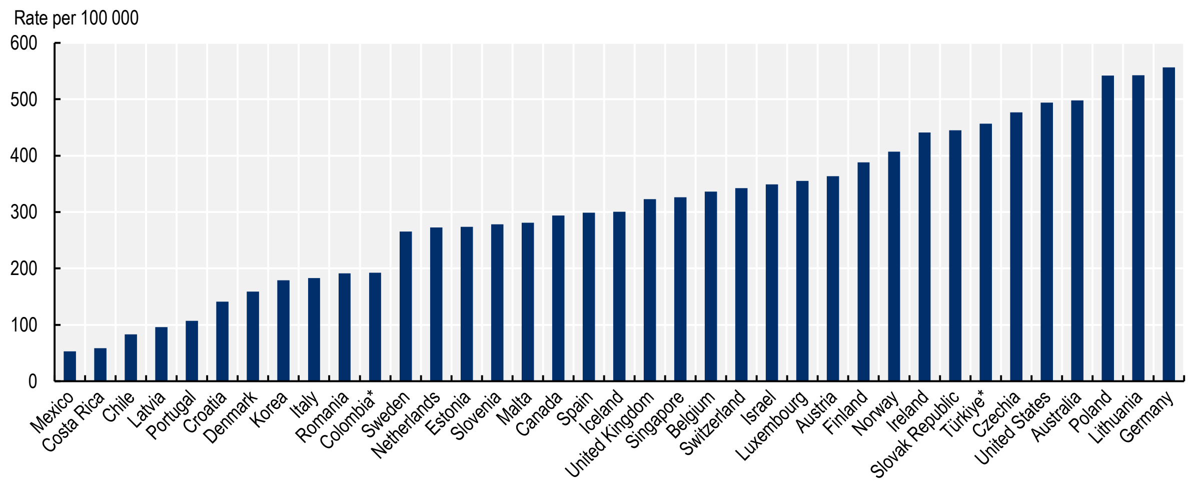
Although the effect of clusters on the outcome variable may be due to health system features that are not captured or controlled for in the analysis or to social determinants of health, results suggest that a primary care oriented system helps improve population health by reducing acute deterioration in people living with asthma, chronic obstructive pulmonary disease, or congestive heart failure. These are widely prevalent long-term conditions that may result in access to hospitals that should not occur in the presence of timely and effective healthcare.

Annex 4.A. Clustering and sensitivity analyses

Output variable

The output variable of this set of analyses is the age-sex standardised rate of avoidable hospital admission rates for asthma, chronic obstructive pulmonary disease (COPD) and congestive heart failure per 100 000 population (Annex Figure 4.A.1)

Annex Figure 4.A.1. Avoidable hospital admission for asthma, chronic obstructive pulmonary disease and congestive heart failure by country, 2022 or latest available year



Source: OECD Health Statistics, January 2024.

Clustering

Three indicators were used to describe characteristics of primary care that could be influenced by actionable policy levers: gatekeeping; continuity of care; financial incentives for primary care physicians to improve quality of care (Annex Table 4.A.1).

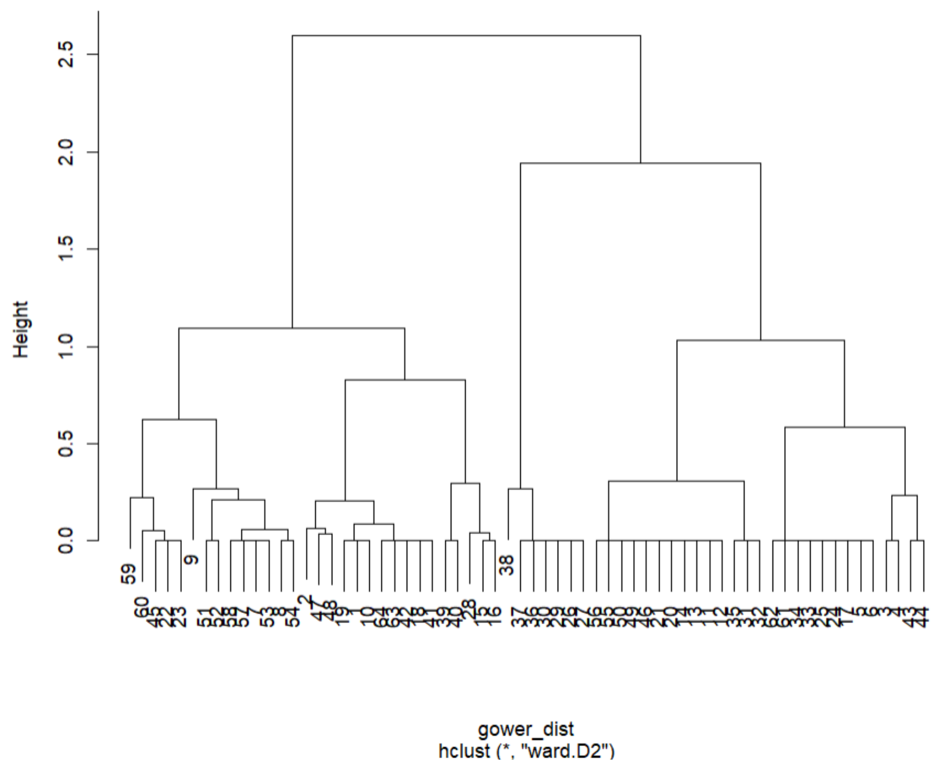
Ward's method was employed to group countries into clusters based on three indicators. Based on Silhouette scores and the visual inspection of the dendrogram, five clusters were identified to best represent groups of health systems (Annex Figure 4.A.2).

Annex Table 4.A.1. Score of indicators by cluster

Indicator	Average for numeric variables and dominance in categorical variables				
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Continuity	C	Mostly C	Mostly B	Mostly C	A
Gatekeeping	C	C	Mostly C	Mostly B	A
Financial incentives to increase healthcare quality in primary healthcare	2.1	0	2.3	0	0.9

Note: Numeric variables a re-scaled between 0 and 3, with 0 being the lowest score and 3 the highest. Role of primary care in the health system (gatekeeping): A: no requirement and no incentive for referral form PHC; B: financial incentives; C: referral form PHC is required. Continuity of care: A: limited part of the population; B: the majority of the population; C: almost the whole population.

Annex Figure 4.A.2. Dendogram



Five clusters were identified as containing elements that were similar among themselves and dissimilar to elements belonging to other groups (Annex Figure 4.A.3).

Annex Figure 4.A.3. Health systems by cluster. Data driven approach

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Australia	Costa Rica	Canada	Austria	Greece
Czechia	Finland	Chile	Belgium	Iceland
Estonia	Ireland	France	Germany	Japan
Hungary	Italy	Poland	Israel	Korea
Latvia	Norway	Portugal	Luxembourg	
Lithuania	Slovenia	Spain	Switzerland	
Netherlands		Sweden		
United Kingdom				

Statistical approach

Pooled ordinary least squares regression

A pooled ordinary least squares regression was used to estimate the relationship between various socio-economic, environmental, and health-related factors and age-sex standardised avoidable hospital admissions (Equation 4). The dependent variables – rate of avoidable admissions – is operationalised as the logarithm to improve model fit. The explanatory variables include Gross Domestic Product (*GDP*) per capita, adjusted for purchasing power parity, which serves as a control for wealth of the country. The unemployment rate was also included in order to capture the economic activity. The Gini index was used to measure income inequality. Educational attainment was represented by the percentage of the population aged 25-65 with tertiary education, which reflects the overall education level of the country. Health-related behaviours were captured through variables such as the percentage of the obese population and the percentage of daily smokers aged 15 and above. Environmental exposure was represented by the mean annual concentration of particulate matter (*PM2.5 in mg per m^3)* per squared kilometre.

In addition to the model variables outlined above, the model included care capacity indicators hospitalisation rates per 100 000 population (*Hospitalisation rate*) and Primary care practitioners (General practitioners, paediatricians and gynaecologist) per 1 000 population (*PCP*).

Additionally, Nonmodifiable health system characteristics ($NonModChar$) were included to account for aspects of the health system that are not subject to change in the short term. Finally, the model integrated the interest dummy variables ($Clusters$) representing the five clusters of countries. The analysis is conducted using the "plm" package in R.

$$\log\left(Rate\ of\ avoidable\ admissions\ per\ 100\ 000\ population_{it}\right) = \alpha + \beta_0 GDP_{it} + \beta_1\ Education_{it} + \beta_2 Beds_{it} + \beta_3 Gini_{it} + \beta_4 Elderly + \beta_5 Obesity_{it} + \beta_6 tobacco_{it} + \beta_7 PCP_{it} + \beta_8 COVID_{it} + \beta_9 NonModChar_i + \beta_{10}\boldsymbol{Clusters_i} + \varepsilon_{it}$$

Equation 4

With i representing OECD countries and t the year in the 2016-22 period.

Assumption testing

Several tests were conducted to assess the underlying assumptions of the panel models. A Chow test was first conducted for poolability using the pooltest function (R package plm), which examines whether the panel data model benefits from individual effects beyond the pooled model. For both models, this test suggested potential individual effects across entities, significant at 5% error. Due to the static nature of the interest variables, they could not be adequately controlled by a fixed effects model at country level without losing these variables from the analysis, thus becoming one of the limitations of the approach. Nevertheless, the standard errors were clustered at the country level to account for potential country level fixed effects and not overestimate the significance of the results. Similarly, the significance was tested, and time fixed effects were included when appropriate by calculating F Test for Individual and/or Time Effects. In the final models of Chapter 3 and 4, time fixed effects were not significant when adding the COVID-19 variable.

Additionally, Pesaran's test of cross-sectional dependence was applied to investigate the independence of errors across panel units. For heteroskedasticity, a Breusch-Pagan test indicated the presence of heteroskedastic error terms in both models, which was addressed by computing robust standard errors using the Arellano method. The Breusch-Godfrey test was conducted for autocorrelation, which checks for serial correlation in the residuals, and assessed multicollinearity among independent variables using the Variance Inflation Factor.

The results of the main model are reported in Annex Table 4.A.2. The variables in the model were removed one by one, with the overall model conclusions remaining strong to the different specifications. The final model corresponds to the one with the highest R squared.

Annex Table 4.A.2. Panel regression results (robust estimates)

Variable	Estimate	SE	P-values
(Intercept)	2.61E+00	4.59E-01	6.49E-07
GDP	-2.54E-06	2.45E-06	3.04E-01
Education	1.91E-02	4.73E-03	1.85E-04
Gini	3.29E+00	1.36E+00	1.88E-02
Obese	2.12E-02	9.19E-03	2.53E-02
tobacco	1.53E-02	1.00E-02	1.34E-01
PM2.5_km	4.20E-04	2.13E-04	5.37E-02
Hospitalization rate	6.16E-05	1.17E-05	2.92E-06
PCP_per1000	-1.59E-01	7.83E-02	4.76E-02
Post_Covid	-1.84E-01	4.67E-02	2.60E-04
NonModChar2 (versus NonModChar1)	1.40E-01	8.17E-02	9.22E-02
NonModChar3 (versus NonModChar1)	7.24E-01	1.83E-01	2.32E-04
NonModChar4 (versus NonModChar1)	-1.42E-01	1.66E-01	3.97E-01
Cluster 2 (versus cluster 1)	5.18E-01	1.06E-01	1.07E-05
Cluster 3 (versus cluster 1)	3.96E-01	1.24E-01	2.44E-03
Cluster 4 (versus cluster 1)	5.71E-01	1.05E-01	1.50E-06
Cluster 5 (versus cluster 1)	5.37E-01	1.12E-01	1.44E-05
R2	0.83		
R2adj	0.78		
F-statistic	*15.27*		

Note: Significant result at *0.05, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.
NonModChar1: Residence-based or single payer + Relatively high spending power in health of subnational decision makers.
NonModChar2: Residence-based or single payer + Relatively low spending power in health of subnational decision makers.
NonModChar3: Multiple insurers + Relatively high spending power in health of subnational decision makers.
NonModChar4: Multiple insurers + Relatively low spending power in health of subnational decision makers.

Sensitivity analyses

A series of sensitivity analyses were performed to test the significance of time fixed effects and the robustness of findings to different outcomes variables and time windows. Tested outcomes were the rates of avoidable admissions by condition (Asthma, COPD, Heart Failure, and a combination of Asthma and COPD), and substituting the hospitalisation rate with the rate of hospital beds per 1 000 population.

- Admission rates for Asthma: the model did not indicate the need for time-fixed effects. Key control variables such as Gini, obesity, hospitalisation rate and smoking rates were positive and significant, while primary care physicians showed a negative correlation with asthma admission rates. Similar to the main model, cluster 4 and cluster 2 had a significant impact, showing a positive relation with admissions compared to cluster 1. On the contrary cluster 5 was associated with lower asthma admissions, but not statistically significant (Annex Table 4.A.3).

Annex Table 4.A.3. Panel regression results using admission rates for Asthma as output variable

Variable	Estimate	SE	P-values
(Intercept)	-8.59E-01	7.06E-01	2.29E-01
GDP	1.34E-05	4.96E-06	9.24E-03
Education	-2.09E-02	6.87E-03	3.67E-03
Gini	6.13E+00	2.20E+00	7.52E-03
Obese	4.61E-02	1.50E-02	3.41E-03
tobacco	5.03E-02	1.26E-02	2.02E-04
PM2.5_km	8.42E-04	3.29E-04	1.35E-02
Hospitalization rate	7.64E-05	1.39E-05	1.24E-06
PCP_per1000	-1.99E-01	8.75E-02	2.74E-02
Post_Covid	-2.10E-01	5.55E-02	4.15E-04
NonModChar2 (versus NonModChar1)	-3.10E-01	1.23E-01	1.50E-02
NonModChar3 (versus NonModChar1)	4.46E-01	3.03E-01	1.48E-01
NonModChar4 (versus NonModChar1)	-7.84E-01	2.44E-01	2.32E-03
Cluster 2 (versus cluster 1)	5.99E-01	1.51E-01	2.27E-04
Cluster 3 (versus cluster 1)	1.26E-01	1.96E-01	5.24E-01
Cluster 4 (versus cluster 1)	5.48E-01	1.78E-01	3.31E-03
Cluster 5 (versus cluster 1)	-2.52E-01	1.98E-01	2.09E-01
R2	0.92		
R2adj	0.89		
F-statistic	*35.35*		

Note: Significant result at *0.05, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.

- Admission rates for COPD: the model revealed significant time-fixed effects. Control variables displayed similar directional behaviour to the main model, but only Gini were significant. The direction and significance of the cluster variables remained robust (Annex Table 4.A.4).

Annex Table 4.A.4. Panel regression results using admission rates for COPD as output variable

Variable	Estimate	SE	P-values
(Intercept)	1.20E-05	4.30E-06	7.87E-03
GDP	1.79E-02	6.29E-03	6.46E-03
Education	1.00E+01	2.21E+00	4.27E-05
Gini	7.53E-03	1.78E-02	6.74E-01
Obese	-1.07E-02	1.34E-02	4.29E-01
tobacco	2.94E-04	2.55E-04	2.55E-01
PM2.5_km	1.13E-05	1.78E-05	5.28E-01
Hospitalization rate	-1.45E-01	1.02E-01	1.65E-01
PCP_per1000	9.82E-02	1.58E-01	5.38E-01
NonModChar2 (versus NonModChar1)	7.47E-01	3.09E-01	1.97E-02
NonModChar3 (versus NonModChar1)	8.68E-02	4.08E-01	8.33E-01
NonModChar4 (versus NonModChar1)	8.41E-01	1.64E-01	5.71E-06
Cluster 2 (versus cluster 1)	5.62E-01	1.88E-01	4.45E-03
Cluster 3 (versus cluster 1)	9.67E-01	2.32E-01	1.35E-04
Cluster 4 (versus cluster 1)	8.80E-01	2.73E-01	2.30E-03
Cluster 5 (versus cluster 1)	1.20E-05	4.30E-06	7.87E-03
R2	0.82		
R2adj	0.74		
F-statistic	*14.40*		

Note: Significant result at *0.05, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.

- Admission rates for Asthma and COPD combined: the model did not require time-fixed effects. While the direction of control variables remained consistent with the main model, their significance shifted, with only Gini and Education remaining significant, while smoking became significant and hospitalisation rate became insignificant. Clusters showed significant effects on admissions with the same direction as the main model (Annex Table 4.A.5).

Annex Table 4.A.5. Panel regression results using admission rates for Asthma and COPD as output variable

Variable	Estimate	SE	P-values
(Intercept)	1.01E-05	4.22E-06	2.40E+00
GDP	3.29E-02	6.93E-03	4.75E+00
Education	1.04E+01	2.39E+00	4.34E+00
Gini	1.69E-02	1.57E-02	1.08E+00
Obese	3.56E-02	1.65E-02	2.15E+00
tobacco	7.32E-04	3.04E-04	2.41E+00
PM2.5_km	-1.90E-05	1.74E-05	-1.09E+00
Hospitalization rate	-1.43E-01	8.47E-02	-1.69E+00
PCP_per1000	1.99E-01	1.32E-01	1.51E+00
Post_Covid	6.30E-01	2.60E-01	2.42E+00
NonModChar2 (versus NonModChar1)	3.41E-01	3.61E-01	9.45E-01
NonModChar3 (versus NonModChar1)	1.17E+00	1.55E-01	7.56E+00
NonModChar4 (versus NonModChar1)	3.37E-01	1.65E-01	2.04E+00
Cluster 2 (versus cluster 1)	7.63E-01	2.34E-01	3.27E+00
Cluster 3 (versus cluster 1)	8.43E-01	2.87E-01	2.93E+00
Cluster 4 (versus cluster 1)	1.01E-05	4.22E-06	2.40E+00
Cluster 5 (versus cluster 1)	3.29E-02	6.93E-03	4.75E+00
R2	0.83		
R2adj	0.76		
F-statistic	*15.37*		

Note: Significant result at *0.05, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.

- Admission rates for Heart Failure: the model did not require time-fixed effects. Moreover, the direction of controls variables remained generally stable compared to the main model. However, only hospitalisation rates and Education were found to be significant were found to be significant (associated with higher avoidable admissions). None of the clusters of characteristics related to the strength of primary care were found significant (Annex Table 4.A.6).

Annex Table 4.A.6. Panel regression results using admission rates for Congestive Heart Failure as output variable

Variable	Estimate	SE	P-values
(Intercept)	4.29E+00	1.02E+00	1.11E-04
GDP	-1.14E-05	7.86E-06	1.53E-01
Education	1.69E-02	7.42E-03	2.74E-02
Gini	-3.93E+00	2.47E+00	1.18E-01
Obese	9.01E-03	1.70E-02	5.99E-01
tobacco	1.41E-02	1.87E-02	4.56E-01
PM2.5_km	1.85E-04	3.92E-04	6.39E-01
Hospitalization rate	1.11E-04	2.51E-05	5.58E-05
PCP_per1000	-5.50E-02	1.19E-01	6.45E-01
Post_Covid	1.00E-01	8.54E-02	2.47E-01
NonModChar2 (versus NonModChar1)	3.44E-02	1.74E-01	8.45E-01
NonModChar3 (versus NonModChar1)	5.36E-01	4.69E-01	2.59E-01
NonModChar4 (versus NonModChar1)	-4.15E-01	3.34E-01	2.20E-01
Cluster 2 (versus cluster 1)	-1.34E-02	2.22E-01	9.52E-01
Cluster 3 (versus cluster 1)	2.92E-01	3.12E-01	3.55E-01
Cluster 4 (versus cluster 1)	1.10E-01	4.09E-01	7.89E-01
Cluster 5 (versus cluster 1)	2.34E-01	3.34E-01	4.87E-01
R2	0.71		
R2adj	0.62		
F-statistic	*7.64*		

Note: Significant result at *0.05, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.

- Using the rate of hospital beds: the direction and significance of the interest variables in the model remains stable compared to the base model that used hospitalisation rates (Annex Table 4.A.7).

Annex Table 4.A.7. Panel regression results using the rate of hospital beds instead of the hospitalisation rate

Variable	Estimate	SE	P-values
(Intercept)	3.68E+00	3.50E-01	2.09E-15
GDP	2.28E-06	2.05E-06	2.71E-01
Education	1.08E-03	3.91E-03	7.83E-01
Gini	3.64E+00	1.02E+00	7.04E-04
Obese	2.48E-02	7.99E-03	2.90E-03
tobacco	-1.20E-02	6.78E-03	8.19E-02
PM2.5_km	-1.74E-04	2.07E-04	4.05E-01
Hospitalization rate	2.78E-01	5.45E-02	3.56E-06
PCP_per1000	-2.80E-01	7.91E-02	7.73E-04
Post_Covid	-2.55E-01	4.50E-02	4.18E-07
NonModChar2 (versus NonModChar1)	-1.21E-01	1.03E-01	2.44E-01
NonModChar3 (versus NonModChar1)	5.93E-01	1.71E-01	9.79E-04
NonModChar4 (versus NonModChar1)	-6.14E-01	2.29E-01	9.37E-03
Cluster 2 (versus cluster 1)	5.45E-01	9.92E-02	7.88E-07
Cluster 3 (versus cluster 1)	3.81E-01	1.02E-01	3.89E-04
Cluster 4 (versus cluster 1)	8.81E-01	1.15E-01	1.77E-10
Cluster 5 (versus cluster 1)	2.63E-01	1.21E-01	3.30E-02
R2	8.06E-01		
R2adj	7.56E-01		
F-statistic	1.61E+01		

Note: Significant result at *0.05, **0.001, ***0.000 level. The model uses the Arellano method for heteroskedasticity-consistent standard errors (White) clustered in time (year) and group (Country). The outcome of the model is log transformed.

Notes

[1] Colombia, Hungary, Mexico, New Zealand, the Slovak Republic and Türkiye are not included in this analysis as there are missing values for the indicators used to cluster health systems.

References

Agency for Healthcare Research and Quality (2002), *Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions*. [4]

Dougherty, S. and L. Phillips (2019), "The spending power of sub-national decision makers across five policy sectors", *OECD Working Papers on Fiscal Federalism*, No. 25, OECD Publishing, Paris, https://doi.org/10.1787/8955021f-en. [16]

Gumas, E. et al. (2024), *Finger on the Pulse: The State of Primary Care in the U.S. and Nine Other Countries*, https://doi.org/10.26099/p3y4-5g38. [11]

Kringos, D. et al. (2013), "The strength of primary care in Europe: an international comparative study", *British Journal of General Practice*, Vol. 63/616, pp. e742-e750, https://doi.org/10.3399/bjgp13x674422. [14]

Kringos, D. et al. (2013), "Europe's strong primary care systems are linked to better population health but also to higher health spending", *Health Affairs*, Vol. 32/4, pp. 86-694, https://doi.org/10.1377/hlthaff.2012.1242. [10]

Lyhne, C. et al. (2022), "Interventions to Prevent Potentially Avoidable Hospitalizations: A Mixed Methods Systematic Review", *Frontiers in Public Health*, Vol. 10, https://doi.org/10.3389/fpubh.2022.898359. [7]

Macinko, J., B. Starfield and L. Shi (2003), "The Contribution of Primary Care Systems to Health Outcomes within Organization for Economic Cooperation and Development (OECD) Countries, 1970–1998", *Health Services Research*, Vol. 38/3, pp. 831-865, https://doi.org/10.1111/1475-6773.00149. [12]

O'Cathain, A. et al. (2013), "Hospital characteristics affecting potentially avoidable emergency admissions: National ecological study", *Health Services Management Research*, Vol. 26/4, pp. 110-118, https://doi.org/10.1177/0951484814525357. [8]

OECD (2020), *Realising the Potential of Primary Health Care*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/a92adee4-en. [3]

OECD/European Commission (2024), *Health at a Glance: Europe 2024: State of Health in the EU Cycle*, OECD Publishing, Paris, https://doi.org/10.1787/b3704e14-en. [2]

Paris, V., M. Devaux and L. Wei (2010), "Health Systems Institutional Characteristics: A Survey of 29 OECD Countries", *OECD Health Working Papers*, No. 50, OECD Publishing, Paris, https://doi.org/10.1787/5kmfxfq9qbnr-en. [15]

Stange, K., W. Miller and R. Etz (2023), "The Role of Primary Care in Improving Population Health", *The Milbank Quarterly*, Vol. 101/S1, pp. 795-840. [1]

Starfield, B., L. Shi and J. Macinko (2005), "Contribution of Primary Care to Health Systems and Health", *The Milbank Quarterly*, Vol. 83/3, pp. 457-502, https://doi.org/10.1111/j.1468-0009.2005.00409.x. [13]

Van Loenen et al. (2016), "The impact of primary care organization on avoidable hospital admissions for diabetes in 23 countries", https://doi.org/10.3109/02813432.2015.1132883. [6]

Van Loenen, T. et al. (2015), "Organizational aspects of primary care related to avoidable hospitalization: a systematic review", https://doi.org/10.1093/fampra/cmu053. [5]

World Health Organization and the United Nations Children's Fund (UNICEF) (2022), *Primary health care measurement framework and indicators: monitoring health systems through a primary health care lens*, https://iris.who.int/handle/10665/352205. [9]

Annex A. The Health Systems Characteristics survey

The OECD survey on Health Systems Characteristics (HSC) aims to collect key policy and institutional characteristics by which heath care systems can be meaningfully differentiated from one another (Paris, Devaux and Wei, 2010[1]).

The HSC survey provides a structured snapshot of the key characteristics that underpin healthcare insurance and healthcare delivery in a given country. It describes arrangements to organise population coverage, the financing of healthcare insurance and delivery, the organisation of healthcare delivery, focusing on the public/private mix of healthcare provision, provider payment schemes, user choice and competition among providers, as well as the regulation of healthcare supply and prices. The survey also describes key aspects of governance and resource allocation in health systems, such as decentralisation in decision-making and the nature of budget constraints.

The Secretariat collected information on health system characteristics through surveys carried out in 2008, 2012, 2016, and 2023. The 2008 survey included 81 questions, often with multiple items and sub-questions for further details. The 2012, 2016, and 2023 surveys comprised 91, 78, and 77 questions, respectively.

The survey was filled in by 29 OECD member states in 2008. For the 2012 round, the then four new OECD member countries – Chile, Estonia, Israel and Slovenia – filled in the survey too, for a total of 33 countries. 28 OECD member countries and Costa Rica and South Africa filled in the survey in 2016 and 32 OECD countries and Bulgaria replied to the survey in 2023. Twenty-one Latin American and Caribbean countries filled in the survey in 2018 (Lorenzoni et al., 2019[2]).

Countries reviewed responses to the different rounds of the OECD surveys on health system characteristics to also verify the consistency of responses over time. Furthermore, the main characteristics of the health systems of OECD countries are publicly available in two datasets on the OECD data platform (https://data-explorer.oecd.org/). These datasets consist of results for the 2023 round as well as results for the 2016 and 2012 rounds of the survey and can be searched according to country, version of the survey, and question. The data can be displayed online or downloaded in an excel format file.

The raw responses to the survey are of interest for certain issues, but for the purpose of providing input to performance analyses, a more manageable set of indicators is necessary. A set of health system characteristics indicators that can help identify cross-country differences in health system performance was developed, informed by discussion with countries (Table A A.1) (de la Maisonneuve et al., 2016[3]; Lorenzoni et al., 2018[4]; Joumard, André and Nicq, 2010[5]). Additional data sources, such as the System of Health Accounts (SHA), were used to complement responses as appropriate, with the implicit assumption that these captured well a policy dimension.

Table A A.1. List of core indicators by domain

Domain	Indicator	Description
Health financing and coverage arrangements	Degree of user choice for basic coverage	A: residence-based; B: insurance-based, single insurer; C: insurance-based, multiple insurers without choice; D: insurance-based, multiple insurers with choice
	Level of financial protection for healthcare users	Share of healthcare spending financed by household out-of-pocket spending in total health spending
	Patient choice among providers	Whether individuals are free to choose any primary care physician, specialist, or hospital to seek care, face incentives to choose a specific primary care physician, specialist, or hospital, or have limited choice. A higher score means that individuals have more options when selecting their primary care physician, specialist, or hospital
	Role of primary care in the health system (gatekeeping)	Categorical (ordered). Financial incentives or referral obligation to access specialist care. A indicates no requirement and no incentive for referral form PHC; B indicates financial incentives and C indicates that referral form PHC is required.
	"Over the basic" coverage	Role played by private health insurance offering complementary, supplementary or duplicative coverage on a voluntary basis. Based on the share of private voluntary health insurance spending in total health expenditure.
	Continuity of care	It captures the share of the population that has a regular doctor to go for care. A: limited part of the population; B: the majority of the population; C: almost the whole population
Healthcare delivery systems	Incentives for volume increase in physicians' payment methods	Evaluates the predominant mode of payment for physicians in primary care, community, and hospitals (i.e. fee-for-service, capitation, salary), as a proxy for incentives to generate volumes of services. A higher value of the indicators means a larger use of fee-for-service
	Incentives for volume increase in hospitals' payment methods	Predominant mode of payment of hospitals. The higher the score the stronger the incentive to generate volumes. A higher value of the indicators means a larger use of fee-for-service.
	Regulation of medical staff in hospitals	Reflects conditions for recruitment and remuneration of medical staff in hospitals. A: Not regulated; B: Partial regulation; C: Highly regulated.
	Financial incentives to improve quality of care	A higher score reflects a system with stronger incentives to improve quality for primary care physicians, specialists and hospitals.
	Degree of private provision of primary care and outpatient specialist services	Evaluates the degree of private provision in primary care and outpatient specialist care. A higher value of the indicators means that the predominant provision of primary and outpatient specialist care is private.
	Advanced roles of nurses	Shows whether nurses independently provide immunisation, health promotion, routine checks for chronically ill patients and minor procedures. A greater number of positive responses – each adding a score of 1 – triggers a higher score that indicates greater responsibilities for nurses.
	Financial incentives for primary care physicians to improve quality of care	A higher score reflects a system with stronger incentives to improve quality of care for primary care physicians.
Governance and resource allocation	Definition of the health benefit basket	Describes how the benefits covered by basic primary health insurance are defined for medical procedures, pharmaceuticals, and implantable devices. A higher score reflects the definition of a benefit basket at central level by a positive list.
	Use of Health Technology Assessment	Existence and use of health technology structure and capacity to determine benefit coverage, reimbursement level/prices and clinical guidelines. A: Little use; B: Somehow used; C: Highly used.
	Regulation of prices/fees for primary care physicians' services, specialists' services and hospitals' services paid by third-party payers	A score is assigned by adding points for the degree of regulation by institutions financing of basic primary coverage. The indicator considers the average of the scores for physicians' services, specialists' services and hospitals' services paid by third-party payers. A higher score indicates a higher degree of regulation.
	Use of electronic health records	Indicates the extent to which primary care physicians use electronic health records by evaluating eight functions: making appointments, keeping records of consultations, ordering laboratory tests, issuing drug prescriptions, sending drug prescriptions to a pharmacy, sending referral letters to medical specialists, receiving alerts or prompts about a potential problem with drug dose or drug interaction, storing diagnostic test results. A greater number of positive responses – each adding a score of 1 – triggers a higher score that indicates greater information continuity.

Qualitative responses provided to the questionnaire were transformed into quantitative indicators using the same scoring system for the different rounds of the survey. Scores were assigned to reflect the country variance rather than distributing countries evenly across categories (Lorenzoni et al., 2018[4]).

For the indicator "financial incentives to improve quality of care" (Table A A.2), the highest score was assigned when bonuses for physicians – both primary care and specialists – related to preventive care, management of chronic diseases, patient satisfaction, population risk factors, uptake of IT services and data quality and linkage were in place, and bonuses for hospitals related to clinical outcomes, use of appropriate processes, patient satisfaction, patient experience and uptake of IT services were at play. The lowest score was assigned when there were no bonuses for physicians and hospitals. The final score is the simple average of the three providers' scores.

Table A A.2. Scoring system for financial incentives to improve quality of care

Primary care physicians	Score
No incentives	0
Yes, incentives related to:	3
Preventive care	(+0.6)
Management of chronic disease	(+0.6)
Patient satisfaction	(+0.6)
Population risk factors	(+0.6)
Uptake of IT services	(+0.6)
Specialists	**Score**
No incentives	0
Yes, incentives related to:	3
Preventive care	(+0.6)
Management of chronic disease	(+0.6)
Patient satisfaction	(+0.6)
Uptake of IT services	(+0.6)
Data quality and linkage	(+0.6)
Hospitals	**Score**
No incentives	0
Yes, incentives related to:	3
Clinical outcomes	(+0.6)
Use of appropriate processes	(+0.6)
Patient satisfaction	(+0.6)
Patient experience	(+0.6)
Uptake of IT services	(+0.6)

In this analysis, a health system was qualified as reporting "Weak" financial incentives for quality if the score was < 3, "Limited" incentives if the score was => 3 and < 4, and "Strong" incentives if the score was = > 4.

For the indicator "financial incentives for primary care physicians to improve quality of care" (Table A A.3), the highest score was assigned when bonuses for physicians – both primary care and specialists – related to preventive care, management of chronic diseases, patient satisfaction, population risk factors, uptake of IT services and data quality and linkage were in place The lowest score was assigned when there were no bonuses for physicians. The final score is the simple average of the scores for the two types of providers.

Table A A.3. Scoring system for financial incentives for primary care physicians to improve quality of care

Primary care physicians	Score
No incentives	0
Yes, incentives related to:	3
Preventive care	(+0.6)
Management of chronic disease	(+0.6)
Patient satisfaction	(+0.6)
Population risk factors	(+0.6)
Uptake of IT services	(+0.6)
Specialists	**Score**
No incentives	0
Yes, incentives related to:	3
Preventive care	(+0.6)
Management of chronic disease	(+0.6)
Patient satisfaction	(+0.6)
Uptake of IT services	(+0.6)
Data quality and linkage	(+0.6)

In this analysis, a health system was qualified as reporting “Limited” financial incentives to physicians for quality if the score was < 3, and “Strong” incentives if the score was = > 3.

The indicator “Degree of private provision of primary care and outpatient specialist services” reports information on private provision of primary care and outpatient specialists’ services. The score is computed as the average of the two sub-scores: private provision of primary care services and private provision of out-patient specialists’ services. The highest score was assigned when the predominant provision of primary care and out-patient specialist services was private only (Table A A.4).

Table A A.4. Scoring system for the degree of private provision of primary care and outpatient specialist services

	Predominant mode	Second mode	Score
Provision of primary care services			
Public healthcare centres only	Public centres		0
Mixed of public and private provision	Public centres	Private clinics, private groups/ solo practice	3
	Private/ solo practice	Public centres	
Private provision only	Private clinics, private groups/ solo practice		6
	Private clinics, private groups/ solo practice	Private clinics, private groups/ solo practice	
Provision of out-patient specialist services			
Public healthcare centres only	Public centres, public hospitals		0
	Public centres, public hospitals	Public centres, public hospitals	
Mixed of public and private provision	Public centres, public hospitals	Private clinics, private groups/ solo practice, private hospitals	3
	Private clinics, private groups/ solo practice, private hospitals	Public centres, public hospitals	
Private provision only	Private clinics, private groups/ solo practice, private hospitals		6
	Private clinics, private groups/ solo practice, private hospitals	Private clinics, private groups/ solo practice, private hospitals	

The indicator “patient choice among providers” is based on responses to questions asking whether individuals are free to choose any doctor or hospital, face incentives to choose a specific doctor or hospital, or have a limited choice. The highest score was assigned when individuals had the possibility to choose any provider, whereas the lowest score was assigned when individuals didn’t have the freedom to choose a provider (Table A A.5).

Table A A.5. Scoring system for patient choice among providers

Provider	Degree of choice	Score (additive)
Primary care physician	Free choice	2
	Incentives to choose	1
	Limited choice	0
Specialist	Free choice	2
	Incentives to choose	1
	Limited choice	0
Hospital	Free choice	2
	Incentives to choose	1.33
	Limited choice but with exceptions (e.g. waiting times)	0.67
	Limited choice	0

The indicator “incentives for volume increase in physicians’ payment methods” is based on responses to questions on primary care physicians and specialists’ predominant mode of payment. A score is assigned according to physicians’ and specialists’ incentives to generate volumes of services, from 0 (salary) to 6 (fee-for-services) (Table A A.6). The final score is a simple average of the two sub-scores for primary care physicians and for specialists (which in turn is the mean of outpatient and inpatient predominant mode of payment for specialists). The higher the score the stronger the incentives to generate volume.

Table A A.6. Scoring system for incentives for volume increase in physicians’ payment methods

Predominant mode of payment of primary care physicians	Score
Salary	0
Mix of salary and capitation	1
Capitation	2
Mix of fee-for-service and salary	3
Mix of fee-for-service, salary and capitation	4
Mix of fee-for-service and capitation	5
Fee-for-service	6
Predominant mode of payment of specialists	Score
Salary	0
Mix of fee-for-service and salary	3
Fee-for-service	6

In this analysis, a health system was qualified as reporting “Limited” use of fee-for-service if the score was < 3, and “Large” use of fee-for-service if the score was = > 3.

The indicator “incentives for volume increase in hospitals’ payment methods” is based on responses to questions on hospitals’ predominant mode of payment. The score is assigned based on the likely impact of a payment method on volumes of care – the higher the incentive to increase volumes the higher the score (Table A A.7). The higher the score the stronger the incentive to generate volume.

Table A A.7. Scoring system for incentives for volume increase in hospitals' payment methods

Predominant mode of payment	Score
Line-item budgets	0
Per diem	1.5
Prospective global budget	3
Payment per case (DRG-like)	4.5
Payment per procedure or service	6

The indicator "definition of the health benefit basket" is based on how the benefits covered by primary health coverage are defined for medical procedures and for pharmaceuticals. The highest score was assigned when the benefit basket was defined at central level by a positive list. If the benefit basket was not considered as the preferred tool for priority setting (benefit basket is not defined), the lowest score was assigned for the indicator (Table A A.8). The final score is the simple average of the scores for medical procedures and pharmaceuticals.

Table A A.8. Scoring system for the definition of the health benefit basket

Medical procedures	Score
Benefit basket is not defined	0
Benefit baskets are defined at local level by providers under budget constraints	1.2
Benefit baskets defined by health insurance funds by a negative list	2.4
Benefit baskets defined by health insurance funds by a positive list	3.6
Benefit baskets are defined at central level by health insurance funds by a negative list	4.8
Benefit baskets are defined at central level by health insurance funds by a positive list	6
Pharmaceuticals	**Score**
Benefit basket is not defined	0
Benefit baskets are defined at local level by providers under budget constraints	1.2
Benefit baskets defined by health insurance funds by a negative list	2.4
Benefit baskets defined by health insurance funds by a positive list	3.6
Benefit baskets are defined at central level by health insurance funds by a negative list	4.8
Benefit baskets are defined at central level by health insurance funds by a positive list	6

The indicator "regulation of prices/fees paid by third-party payers" captures the degree of price/fee regulation for primary care physicians' and hospitals' services. The higher the degree of prices/fees regulation the higher the score assigned (Table A A.9). The final score is the simple average of the scores for primary care physicians and hospitals.

Table A A.9. Scoring system for the regulation of prices/fees paid by third-party payers

Primary care physicians' services	
Third-party payers are "price takers"	0
Prices/fees are negotiated between individual third-party payers and providers	1
Prices/fees are negotiated between collective third-party payers and providers at central or local level, but individual insurers have the ability to negotiate prices for some services	2
Prices/fees are negotiated between collective third-party payers and providers at local level	3
Prices/fees are negotiated between collective third-party payers and providers at central level	4
Prices/fees are set unilaterally by the central government or third-party payers	5
Mix of FFS and capitation. Mix of salary and FFS. Mix of salary, FFS and capitation	5.5
Salaries (or capitation formula) are set at central level by the government	6
Hospital services	
Third-party payers are "price takers"	0
Prices are always negotiated by individual third-party payers and providers	2
Most of prices are set or negotiated at central or local level, but some prices can be negotiated between individual purchasers and providers	3
Prices/rates are negotiated by collective third-party payers and providers at the local level	4
Prices/rates are negotiated by collective third-party payers and providers at the central level	5
Prices/rates or budgets are set unilaterally by the government	6

For the indicator "use of electronic health records", the highest score was assigned when electronic health records were used for evaluating eight functions: making appointments, keeping records of consultations, ordering laboratory tests, issuing drug prescriptions, sending drug prescriptions to a pharmacy, sending referral letters to medical specialists, receiving alerts or prompts about a potential problem with drug dose or drug interaction, storing diagnostic test results (Table A A.10). The lowest score was assigned when electronic health records were not used.

Table A A.10. Scoring system for electronic health records

Primary care physicians	Score
No use of electronic health records	0
Yes, electronic health records used for:	1
Making appointments	(+ 1)
Keeping records of consultations	(+ 1)
Ordering laboratory tests	(+ 1)
Issuing prescription drugs	(+ 1)
Sending drugs prescriptions to a pharmacy	(+ 1)
Sending referral letters to medical specialists	(+ 1)
Receiving alerts or prompts about a potential problem with drug dose or drug interaction	(+ 1)
Storing diagnostic test results	(+ 1)

Table A A.11 to Table A A.13 report the score by indicator by country for two rounds of the OECD survey on health systems characteristics.

Table A A.11. Score by indicator by country, 2016 and 2023

Country	Year	Degree of user choice for basic coverage	Level of financial protection for healthcare users (%)	Patient choice among providers (6 = largest choice)	Role of primary care in the health system (gatekeeping)	"Over the basic" coverage (%)
Australia	2016	A	16.9	5.3	C	14.3
Australia	2023	A	14.9	4.3	C	13.2
Austria	2016	C	19.2	5.3	A	6.8
Austria	2023	C	15.8	5.3	A	6.4
Belgium	2016	D	18.5	6.0	B	4.7
Belgium	2023	D	17.9	6.0	B	4.5
Canada	2016	A	15.6	2.7	C	14.5
Canada	2023	A	14.9	2.7	C	13.8
Chile	2016	D	34.8	2.7	C	6.8
Chile	2018	D	33.2	2.7	C	6.4
Chile	2023		29.8			8.0
Colombia	2016		15.6			8.3
Colombia	2018	D	15.1	0.0	C	7.6
Colombia	2023		13.7			8.0
Costa Rica	2016	B	22.0	0.0	C	2.9
Costa Rica	2018	B	22.4	0.0	C	3.8
Costa Rica	2023	B	20.7	0.0	C	4.8
Czechia	2016	D	14.8	6.0	A	0.9
Czechia	2023	D	12.7	5.0	A	0.9
Denmark	2016	A	13.5	4.0	B	2.3
Denmark	2023		12.9			2.3
Estonia	2016	B	23.7	4.3	C	1.6
Estonia	2023	B	22.1	4.3	C	2.0
Finland	2016	A	19.3	0.7	C	4.5
Finland	2023	A	16.1	0.7	C	4.1
France	2016	C	9.6	5.0	B	7.4
France	2023	C	8.9	5.0	B	6.3
Germany	2016	D	12.8	5.3	B	2.8
Germany	2023	D	11.0	5.3	B	2.5
Greece	2016	B	34.3	3.3	A	4.1
Greece	2023	B	33.3	2.7	A	4.4
Hungary	2016		27.7			4.2
Hungary	2023	B	27.6	6.0	A	2.9
Iceland	2016	A	16.9	4.0	A	1.6
Iceland	2023	A	14.9	4.0	A	1.8
Ireland	2016	A	13.1	6.0	C	14.4
Ireland	2023	A	10.7	6.0	C	12.0
Israel	2016	D	22.8	4.7	B	12.5
Israel	2023	D	19.8	0.7	B	11.0
Italy	2016	A	23.3	4.0	C	2.3
Italy	2023		21.4			2.7
Japan	2016	C	12.9	6.0	A	3.1
Japan	2023	C	11.1	6.0	A	3.2
Korea	2016	B	34.1			8.1
Korea	2023	B	28.0	6.0	A	9.3
Latvia	2016	A	43.3	3.3	B	0.8
Latvia	2023	A	27.0	3.3	B	3.6
Lithuania	2016	B	32.3	6.0	C	1.1

Country	Year	Degree of user choice for basic coverage	Level of financial protection for healthcare users (%)	Patient choice among providers (6 = largest choice)	Role of primary care in the health system (gatekeeping)	"Over the basic" coverage (%)
Lithuania	2023	B	30.2	6.0	C	1.3
Luxembourg	2016	B	10.7	6.0	A	4.3
Luxembourg	2023	B	8.8	6.0	A	4.2
Mexico	2016	C	41.4	0.0	B	7.3
Mexico	2018	C	42.3	0.7	C	8.1
Mexico	2023		41.4			8.4
Netherlands	2016	D	11.3	4.3	C	7.5
Netherlands	2023	D	9.8	4.3	C	5.7
New Zealand	2016		13.6			7.8
New Zealand	2023		12.9			7.9
Norway	2016	A	14.3	6.0	C	0.4
Norway	2023	A	14.1	6.0	C	0.4
Poland	2016	B	22.8	6.0	C	7.9
Poland	2023	B	18.0	6.0	C	7.1
Portugal	2016	A	29.4	0.7	C	8.9
Portugal	2023	A	28.6	0.7	C	8.0
Slovak Republic	2016		18.2			1.4
Slovak Republic	2023	C	19.4	2.3	C	0.9
Slovenia	2016	B	12.0	6.0	C	15.3
Slovenia	2023	B	12.7	6.0	C	13.0
Spain	2016	A	22.0	0.7	C	6.5
Spain	2023	A	21.0	0.7	C	7.4
Sweden	2016	A	14.5	6.0	B	1.2
Sweden	2023	A	13.1	6.0	B	1.0
Switzerland	2016	D	23.0	3.3	B	9.1
Switzerland	2023	D	22.3	3.3	B	8.6
Türkiye	2016	B	16.5	3.3	A	5.1
Türkiye	2023	B	16.3			4.9
United Kingdom	2016	A	14.5	4.0	C	5.0
United Kingdom	2023	A	13.9	4.0	C	4.5
United States	2016		11.7			5.5
United States	2023	D	10.7	0.0	B	5.7

Note: The indicators "Level of financial protection for healthcare users" and "Over the basic" coverage use the latest measure available in the OECD Data explorer. Degree of user choice for basic coverage: A: residence-based; B: insurance-based, single insurer; C: insurance-based, multiple insurers without choice; D: insurance-based, multiple insurers with choice. Role of primary care in the health system (gatekeeping): A: no requirement and no incentive for referral form PHC ("Weak" gatekeeping); B: financial incentives ("Medium"); C: referral form PHC is required ("Strong").
Source: Health System Characteristics Surveys; OECD Health Statistics.

Table A A.12. Score by indicator by country, 2016 and 2023 (continued)

Country	Year	Incentives for volume increase in physicians' payment methods (6 = largest incentives)	Incentives for volume increase in hospitals' payment methods (6 = largest incentives)	Regulation of medical staff in hospitals	Financial incentives to improve quality of care (6 = largest incentives)	Degree of private provision of primary care and outpatient specialist services (6 = largest private provision)	Definition of the health benefit basket (6 = defined at central level through a positive list)	Use of Health Technology Assessment	Regulation of prices/fees paid by third-party payers (6 = strongest degree of regulation)
Australia	2016		4.5	A	1.6	6.0	5.2	C	5.5
Australia	2023	4.0	4.5	A	0.8		5.2	C	5.5
Austria	2016	2.7	4.5	B	0.0	4.5	5.7	C	4.2
Austria	2023	2.7	4.5	B	0.0		5.7	C	4.2
Belgium	2016	6.0	4.5		0.0	6.0	6.0	C	5.0
Belgium	2023	6.0	4.5		0.0		6.0	C	5.0
Canada	2016	5.3	3.0		1.6	4.5	5.0	C	4.0
Canada	2023	5.3	3.0		1.8		4.7	C	4.0
Chile	2016	0.0	3.0	C	4.6	3.0	5.3	C	4.9
Chile	2018	0.0	3.0	C	4.6		5.3	C	4.9
Chile	2023								4.9
Colombia	2016								
Colombia	2018	1.0	0.0	A	0.0	3.0	6.0	B	4.3
Colombia	2023								
Costa Rica	2016	0.0	0.0	C	0.0			C	6.0
Costa Rica	2018	0.0	0.0	C	0.0	1.5		C	
Costa Rica	2023	0.0	0.0	C	0.0		6.0	C	6.0
Czechia	2016	3.7	4.5	B	1.6	6.0		B	4.7
Czechia	2023	3.7	4.5	B	4.2				4.7
Denmark	2016	3.0	3.0	A	0.6	4.5	2.3	C	4.7
Denmark	2023								
Estonia	2016			B	1.4	3.0	6.0	C	5.3
Estonia	2023	1.7	6.0	B	1.6		6.0	C	5.3
Finland	2016	0.0	4.5	C	0.0	0.0	3.0	C	6.0
Finland	2023	0.0	4.5	B	0.0		3.0	C	6.0
France	2016	3.3	4.5	B	4.6	6.0	6.0	C	5.0
France	2023	3.3	4.5	B	4.0		6.0	C	5.0
Germany	2016		4.5		0.0	6.0	4.7	C	2.7
Germany	2023	4.0	4.5	A	0.0		4.8	C	4.0
Greece	2016	2.7	4.5	C	0.0	4.5	4.7	C	3.7
Greece	2023	2.7	4.5	C	0.0		4.7	C	3.7
Hungary	2016					1.5			
Hungary	2023	0.3	4.5	B	1.4		6.0	C	5.5
Iceland	2016	2.0	3.0	B	0.0	3.0	4.9	C	4.7
Iceland	2023	2.0	4.5	B	0.0		4.9	C	4.7
Ireland	2016		3.0			4.5	3.0	C	
Ireland	2023		3.0		0.0		4.7	C	
Israel	2016	0.0	0.0	B	0.0	1.5	6.0	C	1.0

Country	Year	Incentives for volume increase in physicians' payment methods (6 = largest incentives)	Incentives for volume increase in hospitals' payment methods (6 = largest incentives)	Regulation of medical staff in hospitals	Financial incentives to improve quality of care (6 = largest incentives)	Degree of private provision of primary care and outpatient specialist services (6 = largest private provision)	Definition of the health benefit basket (6 = defined at central level through a positive list)	Use of Health Technology Assessment	Regulation of prices/fees paid by third-party payers (6 = strongest degree of regulation)
Israel	2023	1.3	0.0	B	1.2		6.0	C	1.0
Italy	2016	0.7	4.5	B	0.0	0.0	5.5	C	
Italy	2023								
Japan	2016		0.0			6.0	6.0	C	4.3
Japan	2023		0.0				6.0	C	4.3
Korea	2016					6.0			
Korea	2023	4.0	6.0	B	4.0		6.0	C	4.3
Latvia	2016	3.7	3.0	B	1.4	3.0	4.9	C	5.3
Latvia	2023	3.7	3.0	B	1.4		6.0	C	5.3
Lithuania	2016		4.5		1.4	3.0		C	5.3
Lithuania	2023	0.7	4.5	B	1.4		6.0	C	5.3
Luxembourg	2016	6.0	3.0	A	0.4	6.0	6.0	C	4.7
Luxembourg	2023	6.0	3.0	A	0.0		6.0	C	4.7
Mexico	2016		3.0	B	1.6	3.0	3.0	C	
Mexico	2018	0.0	3.0	B	0.0		6.0	B	5.0
Mexico	2023								
Netherlands	2016	5.7		B	3.0	6.0	3.0	C	2.3
Netherlands	2023	4.7		B	3.2		4.0	C	2.3
New Zealand	2016					6.0			
New Zealand	2023								
Norway	2016	1.7	3.0	B	2.0	3.0	5.5	C	5.0
Norway	2023	2.7	3.0	B	0.0		5.5	C	5.0
Poland	2016		4.5	A	3.4	0.0	6.0	C	5.5
Poland	2023	2.7	4.5	A	3.2		6.0	C	5.5
Portugal	2016	1.0	3.0	C		0.0	4.0	B	4.5
Portugal	2023	2.0	3.0	B	5.0		4.0	B	6.0
Slovak Republic	2016					4.5			3.0
Slovak Republic	2023	2.7	3.0	B	4.2		6.0	C	3.0
Slovenia	2016	0.0	4.5	B	0.0	1.5	6.0	C	4.3
Slovenia	2023	0.0	4.5	B	0.0		6.0	C	4.3
Spain	2016	1.3	3.0	B	4.2	3.0	6.0	C	4.2
Spain	2023	1.3	3.0	B	5.0		6.0	C	4.2
Sweden	2016	0.0	3.0	A	3.8	3.0	3.0	C	4.8
Sweden	2023	0.0	3.0	A	3.8		3.0	C	4.8
Switzerland	2016	5.0	4.5	A	0.0	4.5	4.9	C	1.7
Switzerland	2023	5.0	4.5	A	0.0		4.9	C	1.7
Türkiye	2016	3.0	3.0	C	0.0	1.5	6.0	C	
Türkiye	2023						6.0		
United Kingdom	2016	1.3	4.5	B	5.0	0.0	4.0	C	4.8
United Kingdom	2023	1.3	4.5	B	4.4		4.0	C	4.8

Country	Year	Incentives for volume increase in physicians' payment methods (6 = largest incentives)	Incentives for volume increase in hospitals' payment methods (6 = largest incentives)	Regulation of medical staff in hospitals	Financial incentives to improve quality of care (6 = largest incentives)	Degree of private provision of primary care and outpatient specialist services (6 = largest private provision)	Definition of the health benefit basket (6 = defined at central level through a positive list)	Use of Health Technology Assessment	Regulation of prices/fees paid by third-party payers (6 = strongest degree of regulation)
United States	2016								
United States	2023	3.3		A	0.0		4.0	B	0.0

Note: Regulation of medical staff in hospitals: A: Not regulated; B: Partial regulation; C: Highly regulated. Use of Health Technology Assessment: A: Little use; B: Somehow used; C: Highly used.
Source: Health System Characteristics Surveys.

Table A A.13. Score by indicator by country, 2016 and 2023 (continued)

Country	Year	Use of electronic health records (9 = largest use)	Advanced roles of nurses (4 = most advanced roles)	Financial incentives for primary care physicians to improve quality of care (6 = largest incentives)	Continuity of care
Australia	2016	9		4.8	C
Australia	2023	8	4	4.8	C
Austria	2016	8		0	B
Austria	2023	6		0	B
Belgium	2016	1	0	0	C
Belgium	2023	6	0	0	C
Canada	2016	4	2	4.8	B
Canada	2023	5	2	5.4	B
Chile	2016	8	4	4.8	A
Chile	2018	8	4	4.8	
Colombia	2018	2	3	0	
Costa Rica	2016	9	3	0	C
Costa Rica	2018	9	4	0	
Costa Rica	2023	8	2	0	C
Czechia	2016	2	2	4.8	C
Czechia	2023	4	3	4.8	C
Denmark	2016	8	1	0	
Denmark	2023				
Estonia	2016	9	4	4.2	C
Estonia	2023	8	4	4.8	C
Finland	2016	9	4	0	C
Finland	2023	8	4	0	C
France	2016	1	2	4.8	B

Country	Year	Use of electronic health records (9 = largest use)	Advanced roles of nurses (4 = most advanced roles)	Financial incentives for primary care physicians to improve quality of care (6 = largest incentives)	Continuity of care
France	2023	1	2	4.8	B
Germany	2016			0	C
Germany	2023	6	3	0	C
Greece	2016	9	3	0	A
Greece	2023	8	3	0	A
Hungary	2016				C
Hungary	2023			4.2	C
Iceland	2016	7	1	0	A
Iceland	2023	7	4	0	A
Ireland	2016	7	4	0	B
Ireland	2023	6	4	0	B
Israel	2016	8	4	0	C
Israel	2023	8	4	0	C
Italy	2016	5	0	0	B
Italy	2023				B
Japan	2016				A
Japan	2023	3	1	0	A
Korea	2016				A
Korea	2023	3	0	3.6	A
Latvia	2016	3	4	4.2	C
Latvia	2023	3	2	4.2	C
Lithuania	2016	0	1	4.2	C
Lithuania	2023	7	2	4.2	C
Luxembourg	2016	3	0	0	C
Luxembourg	2023	0	1	0	C
Mexico	2016			4.8	
Mexico	2018	0	2	0	
Mexico	2023				
Netherlands	2016	8	4	3	C
Netherlands	2023	8	4	3.6	C
Norway	2016	7	0	0	C
Norway	2023	8	0	0	C
Poland	2016	1	4	3	B
Poland	2023	6	3	3	B
Portugal	2016	8	4	4.8	B
Portugal	2023	7	4	5.4	B
Slovak Republic	2023	2	1	5.4	
Slovenia	2016	5	0	0	C
Slovenia	2023	5	0	0	C
Spain	2016	9	4	4.8	B

Country	Year	Use of electronic health records (9 = largest use)	Advanced roles of nurses (4 = most advanced roles)	Financial incentives for primary care physicians to improve quality of care (6 = largest incentives)	Continuity of care
Spain	2023	8	4	4.8	B
Sweden	2016	9	4	5.4	B
Sweden	2023	8	4	5.4	B
Switzerland	2016	6	1	0	C
Switzerland	2023	3	1	0	C
Turkiye	2016	6	2	0	
United Kingdom	2016	8	4	4.2	C
United Kingdom	2023	8	4	4.2	C
United States	2016				B
United States	2023	1	4	0	B

Note: Continuity of care: A: limited part of the population has a regular doctor to go for care ("Weak" continuity); B: the majority of the population has a regular doctor to go for care ("medium"); C: almost the whole population has a regular doctor to go for care ("Strong").

References

de la Maisonneuve, C. et al. (2016), "The Role of Policy and Institutions on Health Spending", *Health Economics*, Vol. 26/7, pp. 834-843, https://doi.org/10.1002/hec.3410. [3]

Joumard, I., C. André and C. Nicq (2010), "Health Care Systems: Efficiency and Institutions", *OECD Economics Department Working Papers*, No. 769, OECD Publishing, Paris, https://doi.org/10.1787/5kmfp51f5f9t-en. [5]

Lorenzoni, L. et al. (2018), "Which policies increase value for money in health care?", *OECD Health Working Papers*, No. 104, OECD Publishing, Paris, https://doi.org/10.1787/a46c5b1f-en. [4]

Lorenzoni, L. et al. (2019), "Health systems characteristics: A survey of 21 Latin American and Caribbean countries", *OECD Health Working Papers*, No. 111, OECD Publishing, Paris, https://doi.org/10.1787/0e8da4bd-en. [2]

Paris, V., M. Devaux and L. Wei (2010), "Health Systems Institutional Characteristics: A Survey of 29 OECD Countries", *OECD Health Working Papers*, No. 88, OECD Publishing Paris, https://doi.org/10.1787/5jlz3kbf7pzv-en. [1]

How Do Health System Features Influence Health System Performance?

International comparisons are an important tool for benchmarking health system performance, shedding light on health systems' relative strengths and weaknesses. The present work examines how different groups of countries sharing similar health system characteristics perform relative to others. To make valid and useful comparisons, health systems may be grouped in ways that resonate with policy makers in countries and reflect the policy question at hand. The report specifically addresses three key policy areas: the influence of the overall design of health systems on performance, the role of financial incentives to providers and the role of a strong primary care system. The report shows that there is no indication that any one group of health systems would systematically outperform another. It further provides evidence that there is room for health systems sharing the same broad characteristics to improve performance by borrowing elements from other systems. Rather than engaging in large-scale system reforms, focusing on more targeted policy changes may be a better avenue for improving performance.

PRINT ISBN 978-92-64-82221-4
PDF ISBN 978-92-64-81481-3
9 789264 822214