Title: Identifying a New Typology of European Health Systems: A Data-Driven Clustering Analysis

Authors: Poe Hnin Phyu, Laiba Khan, Maham Khan, Mahmood Ahmad, Nijat Rizayev, Manpreet Kour, Touqeer Rana, Andrew Woo, Parichatra Homhuan.

Affiliation : All contributing authors are Affiliated with Royal Free Hospital, London

Corresponding Author:

Joanne Lac

Email: joanne.lac.20@ucl.ac.uk

Affiliation : University College London (UCL)

Abstract

Traditional health system classifications may not accurately reflect the complex realities of modern European nations. This study takes a data-driven approach to create a new classification. We used WHO data and applied a K-Means clustering algorithm to 52 European countries, focusing on recent figures for doctor density, nurse density, and catastrophic health spending. The analysis revealed three clear clusters: "High-Resource, Low-Burden" systems (n=14), "Physician-Centric, High-Burden" systems (n=18), and "Lower-Resource" systems (n=20). This classification offers a detailed way to understand current health system structures and to inform policy decisions across Europe.

Introduction

Classifying health systems is crucial for public health policy and research. While traditional models like Beveridge and Bismarck have guided analysis , years of reforms have led to hybrid systems, indicating a need for new, evidence-based classifications reflecting current healthcare resources and funding..

Unsupervised machine learning, particularly cluster analysis, offers a robust way to identify such evidence-based groupings without being influenced by historical perspectives. This study aims to use this data-driven method to answer the following research question: What distinct health system profiles exist across Europe based on key indicators of healthcare resources and funding?

Methods

This cross-sectional study used publicly available, aggregated data for 52 countries in the WHO European Region, collected from the WHO's Global Health Observatory and World Health Statistics. The analysis focused on three main indicators: the number of medical doctors (per 10,000 population), the number of nursing and midwifery personnel (per 10,000 population), and the percentage of the population facing catastrophic health expenses (household health spending exceeding 10% of total income).

For each country, the most recent data point for each indicator was selected to create a current profile. Missing values were estimated using a K-Nearest Neighbours method (k=5), a reliable approach that draws on data from similar countries to fill gaps. To ensure that each indicator contributed equally to the analysis, all data were adjusted to have a mean of 0 and a standard deviation of 1.

A K-Means clustering algorithm was applied to divide the 52 countries into distinct groups. Using the elbow method, the optimal number of clusters was identified as three. This method finds the point where adding more clusters yields diminishing returns on model fitting. Ethical approval was not necessary since the study relied solely on public and anonymized country-level data.

Results

The analysis divided the 52 European countries into three distinct clusters. The composition and mean profile of each cluster are detailed in **Table 1**

Cluster 0, labeled "High-Resource, Low-Burden Systems" (n=14), is marked by the highest average density of both doctors and nurses and the lowest rates of catastrophic health expenses. This group includes countries like Germany, Switzerland, and most Nordic nations.

Cluster 1, identified as "Physician-Centric, High-Burden Systems" (n=18), features an above-average number of doctors but a significant shortage of nurses. This cluster struggles with the highest average rate of catastrophic health expenses.

Cluster 2, termed "Lower-Resource Systems" (n=20), has the lowest average density of both doctors and nurses. This larger cluster includes major economies like the United Kingdom.

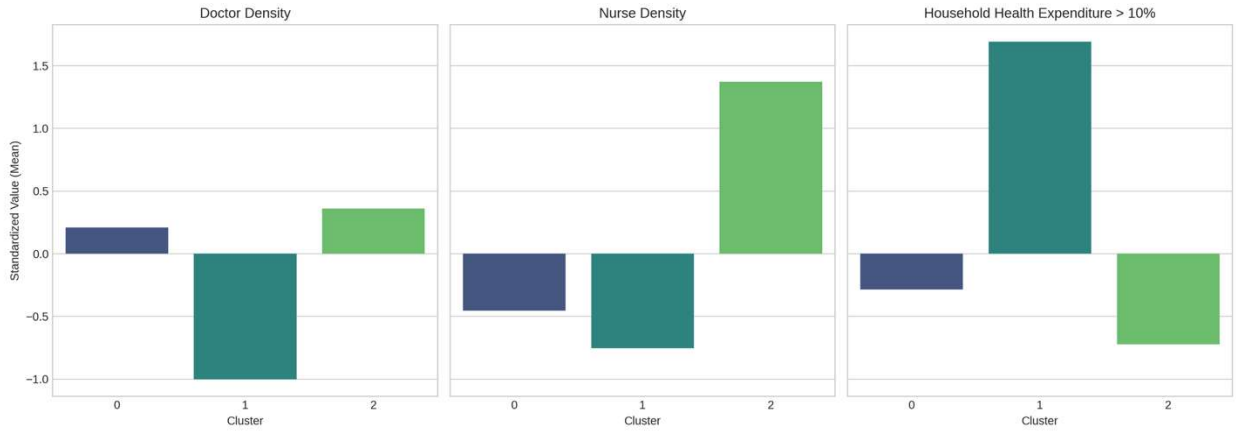
Table 1: Cluster Composition and Mean Indicator Values

Cluster Name	N	Countries	Doctor Density	Nurse Density	Catastrophic

			per 10k	per 10k	Health Expenditure > 10% (%)
0: High-Resource, Low-Burden	14	Austria, Belgium, Czechia, Denmark, Finland, Germany, Iceland, Ireland, Luxembourg, Monaco, Netherlands (Kingdom of the), Norway, Sweden, Switzerland	42.8	155.1	5.5
1: Physician-Centric, High-Burden	18	Andorra, Armenia, Azerbaijan, Bulgaria, Cyprus, Georgia, Greece, Italy, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Malta, Portugal, Republic of	39.7	65.9	19.8

		Moldova, Russian Federation, San Marino, Tajikistan, Ukraine			
2: Lower-Resource	20	Albania, Belarus, Bosnia and Herzegovina, Croatia, Estonia, France, Hungary, Israel, Montenegro, North Macedonia, Poland, Romania, Serbia, Slovakia, Slovenia, Spain, Turkmenistan, Türkiye, United Kingdom, Uzbekistan	25.1	80.5	9.7

Figure 2: Cluster Profiles by Indicator (Standardized Means)



Discussion

This study successfully identified three unique health system profiles in modern Europe, creating a new evidence-based classification that supplements traditional models. The clusters showcase distinct differences in health workforce strategy and financial protection for citizens.

The "High-Resource, Low-Burden" cluster seems to represent a high-investment model of European healthcare, developed by countries from both the Bismarck and Beveridge traditions. In contrast, the "Physician-Centric, High-Burden" cluster closely aligns with the known legacy of the post-Soviet "Semashko" system, which historically favored specialists over nursing and primary care, often resulting in high out-of-pocket expenses.

Classifying the UK and France as "Lower-Resource" systems is a noteworthy finding. This classification reflects contemporary challenges like nursing shortages in the UK's NHS and difficulties in France's primary care workforce. This underscores the importance of data-driven comparisons for evaluating established health systems.

The main limitation of this study is its cross-sectional design, which prevents any causal conclusions. However, by employing a strong imputation method and concentrating on key structural indicators, this analysis offers a valid snapshot of the current European health system landscape.

This typology provides a new basis for policy analysis, enabling countries to identify and learn from structural peers. Future research could use these clusters to explore the factors influencing health system performance and resilience.

Limitations

This study has several limitations. It focused on just three indicators: physician density, nurse density, and catastrophic health expenditure. This choice may have missed important factors like total health expenditure, primary care capacity, and health outcomes. Methodologically, it only used K-means clustering, without comparing it to other methods or checking the stability of the clusters. Furthermore, although the study addressed missing data, it did not provide details on how much data was missing or the methods used for imputation. Sensitivity analyses were also not performed. Because of these reasons, the typology should be seen as preliminary. Future research should include a wider range of indicators and stronger methods..

Funding: No funding was received for this work.

Conflicts of Interest: None declared.

Data Availability Statement: The data used in this article were obtained from the World Health Organization's Global Health Observatory and World Health Statistics repository, which are publicly available. The aggregated data generated and analyzed during this study can be requested from the corresponding author.

Key Points

This study uses machine learning to create a new evidence-based classification of European health systems.

The analysis reveals three distinct models: "High-Resource, Low-Burden," "Physician-Centric, High-Burden," and "Lower-Resource."

The data-driven clusters challenge traditional classifications by reflecting current operational realities.

The findings offer a fresh framework for international comparison and evidence-based health policy.

The classification of countries like the UK and France highlights current workforce challenges in major health systems.

References

1. Ridic G, Gleason S, Ridic O. Comparisons of health care systems in the United States, Germany and Canada. *Mater Sociomed*. 2012;24(2):112-120. https://pmc.ncbi.nlm.nih.gov/articles/PMC3633404/
2. Böhm K, Schmid A, Götze R, Landwehr C, Rothgang H. The 'Bismarck' Health Care System. In: *The Changing Political Economy of the Welfare State*. Springer; 2013.
3. Wendt C, Frisina L, Rothgang H. Healthcare system types: a conceptual framework for comparison. *Soc Policy Adm*. 2009;43(1):70-90. https://doi.org/10.1111/j.1467-9515.2008.00647.x
4. Saltman RB. The theory of health system convergence: A sociological perspective. *Soz Praventivmed*. 2003;48:7-13.
5. Lavecchia AM, Zecchinelli M. A new data-driven methodology to profile and classify healthcare systems. *Health Policy*. 2016;120(5):483-492. https://doi.org/10.1016/j.healthpol.2016.03.006
6. Rechel B, McKee M. Health reform in central and eastern Europe and the former Soviet Union. *Lancet*. 2009;374(9696):1186-1195. https://doi.org/10.1016/S0140-6736(09)61334-9
7. Popovich L, Potapchik E, Shishkin S, Richardson E, Vacarezza M, Mathivet B. The Russian Federation: health system review. *Health Syst Transit*. 2011;13(7):1-190. https://www.ecoi.net/en/file/local/1302492/1930_1421317372_hit-russia-en-web-with-links.pdf
8. Buchan J, Shembavnekar N. *The health and care workforce in England*. The King's Fund; 2022.
9. Chevreul K, Berg Brigham K. France: health system review. *Health Syst Transit*. 2015;17(3):1-210. https://eprints.lse.ac.uk/65775/
10. MacQueen J. Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. 1967:281-297. https://www.cs.cmu.edu/~bhiksha/courses/mlsp.fall2010/class14/macqueen.pdf

Authors:

Mahmood Ahmad – [Royal Free Hospital Nhs ,London]

Poe Hnin Phyu – [Royal Free Hospital Nhs ,London]

Laiba Khan – [Royal Free Hospital Nhs ,London]

Nijat Rizayev – [Royal Free Hospital Nhs ,London]

Parichatra Homhuan – [Royal Free Hospital Nhs ,London]

Manpreet Kour – [Royal Free Hospital Nhs ,London]

Touqeer Rana – [Royal Free Hospital Nhs ,London]

Maham Khan – [Royal Free Hospital Nhs ,London]

Andrew Woo – [Royal Free Hospital Nhs ,London]

Corresponding Author:

Joanne Lac

Email: joanne.lac.20@ucl.ac.uk