ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

ORIGINAL RESEARCH ARTICLE

Forecasting world health expenditures: A hybrid artificial intelligence framework

Taegeon Yu[1], Daipayan Bera[1], Abbas Maazallahi[2], Roschlynn Dsouza[1], Francina Pali[1], Wen-Shan Liu[1], Payam Norouzzadeh[3], Eli Snir[4], and Bahareh Rahmani[1]*

[1]Department of Health and Clinical Outcomes Research, School of Medicine, Saint Louis University, Saint Louis, Missouri, United States of America

[2]Department of Computer Science, School of Mathematics, Statistics, and Computer Science, College of Science, University of Tehran, Tehran, Iran

[3]Department of Analytics, School for Professional Studies, Saint Louis University, Saint Louis, Missouri, United States of America

[4]Department of Business Analytics, Olin Business School, Washington University in Saint Louis, Saint Louis, Missouri, United States of America

Abstract

Global healthcare expenditures continue to rise, posing substantial economic challenges, particularly for low- and middle-income countries (LMICs), where resource constraints intensify the impact. Accurate forecasting, efficient resource allocation, and equitable policy development are essential to address these growing pressures. This study presents a hybrid analytical framework that integrates generative artificial intelligence (AI) with traditional econometric and machine learning models to analyze and predict trends of healthcare expenditure. Utilizing data from the World Bank and World Health Organization, we applied generative adversarial networks, hierarchical clustering, support vector machines, and autoregressive integrated moving average models to uncover spending patterns, simulate policy scenarios, and tackle disparities in health investment. Generative AI played a pivotal role by augmenting sparse and incomplete datasets, particularly from underrepresented LMICs, identifying anomalies, and generating realistic synthetic data to support robust forecasting. This enabled the development of more inclusive, equity-focused health resource planning tools. The results demonstrate improved forecasting accuracy and offer deeper insights into regional and income-based differences in expenditure trends. By combining traditional machine learning with cutting-edge generative models, this study advances a scalable, data-driven approach to support global health decision-making. Ultimately, generative AI is highlighted as a transformative enabler of equitable, informed strategies in the management of global healthcare expenditures.

Keywords: Healthcare expenditure; Generative artificial intelligence; Autoregressive integrated moving average; Health equity; Support vector machines; Low- and middle-income countries

***Corresponding author:**
Bahareh Rahmani
(bahareh.rahmani@health.slu.edu)

Citation: Yu T, Bera D, Maazallahi A, *et al*. Forecasting world health expenditures: A hybrid artificial intelligence framework. *Artif Intell Health*. 2026;3(1):152-163. doi: 10.36922/AIH025170033

Received: April 21, 2025

1st revised: June 22, 2025

2nd revised: July 4, 2025

Accepted: August 12, 2025

Published online: September 22, 2025

1. Introduction

Rising global healthcare expenditure presents a critical challenge for policymakers, especially in low- and middle-income countries (LMICs), where financial and infrastructural limitations exacerbate health inequities. Traditional econometric models often fall short in capturing the complexity of health spending patterns across diverse socioeconomic and regional contexts. Moreover, the scarcity and inconsistency of data in LMICs hinder the development of robust forecasting tools and the formulation of health policies.

Following substantial increases of 7.4% in 2022 and 10.7% in 2023, it is estimated that the growth in global healthcare expenditure will remain high at 9.9% in 2024. While there is debate over whether population aging is the primary cause of rising healthcare costs, many analysts and representatives consider it one of the major contributors.[1] Healthcare expenditure has become increasingly burdensome globally, driven by various factors, including gross domestic product (GDP) growth, aging populations, and healthcare strategies.[2,3] The United States (US) exemplifies the surge in healthcare expenditure across both private and public sectors, with per capita costs and the share of GDP rising sharply due to rapid advancements in medical treatments and technology. Technological advancements, especially in developed nations, such as the US, have significantly expanded treatment options and increased per capita spending, often exceeding expenditures in many countries with universal healthcare structures.[2] Estimates of changes in global health research and development expenditures are crucial for improving and setting boundaries in health research planning.[4,5] However, comparing this information across countries can be a challenging task, as there is a gap in measuring data per area.[4]

Over the past couple of decades, healthcare spending has nearly doubled in many high-income countries, with projections suggesting that global spending could reach $15 trillion within 25 years.[5] In 2021, US healthcare spending was nearly twice that of other high-income countries, with significantly higher costs for inpatient, outpatient, and administrative facilities.[6] In addition, it is estimated that the growth of US per capita spending on health, at an annual average rate of 5.0%, will continue to outpace the GDP growth from 2023 to 2027, stressing the need for systemic reform. At this rate, it is estimated that health spending will reach 17.9% of GDP in 2025, growing slightly faster than the economy. In addition, health spending growth is expected to exceed overall economic growth, reaching 19.7% of GDP within 7 years.[7]

These figures underscore the significance of government regulations and policies in mitigating the upward trend in healthcare costs. However, fluctuations in government revenues, driven by events, such as the COVID-19 pandemic, the Russo-Ukrainian War, inflation, and recessionary pressures, have widened spending disparities between high- and low-income countries. In addition, the weakening global economy has also constrained health funding. The global economic slowdown in 2021 squeezed health budgets, reducing health aid contributions from high-income countries, such as the United Kingdom (UK) and Sweden, thereby diminishing support for low-income nations.[8]

Several interconnected factors contribute to the excessive rise in healthcare expenditures, including expanded insurance coverage, supplier-induced demand, defensive medicine, factor productivity, and technological advances.[9] While these technological developments have improved treatment options and quality of life, they have also contributed to sustained increases in healthcare expenses in the US. Conversely, lifestyle factors, particularly modifiable risk behaviors, such as regular physical activity and healthy dietary choices, can reduce healthcare expenses. Moreover, environmental factors, including carbon dioxide emissions and fossil fuel consumption, negatively impact population health.[10] For example, one study indicates that air pollution in China significantly increases healthcare costs, with the economic burden extending well beyond classic respiratory illnesses.[11]

Innovations in technology have introduced automation and cost-effectiveness into healthcare, transforming research, diagnostics, and treatment delivery.[12] Artificial intelligence (AI) has become a transformative instrument across several fields, particularly in health economics and outcomes research. The application of AI, particularly machine learning, offers novel approaches for enhancing prediction models, economic analysis, and healthcare decision-making trials[13] For example, to evaluate the possibility of patient hospitalization, several machine learning strategies have been developed using prevalent methodologies with insurance claim datasets to improve prediction accuracy.[14,15] Other studies have validated how machine learning is revolutionizing medical commerce.[16,17] It is a crucial task for humanity to enhance the quality of healthcare through machine learning techniques, as it was shockingly found that a high proportion of healthcare expenditure failed to protect countries from COVID-19.[18] Nonetheless, machine learning should be viewed as a complement to, rather than a replacement for, human judgment, allowing human oversight to remain as a core standard.[19]

The integration of AI into healthcare has transformed the landscape of diagnostics, treatment, public health, and health systems management. Over the past decade, advancements in machine learning and deep learning have significantly driven the development of intelligent, data-driven health technologies. For example, Rajkomar *et al.*[20] provided a foundational overview of how machine learning is revolutionizing diagnostics and clinical decision-making, enabling earlier and more accurate treatment strategies. Complementing this, Topol[21] emphasized the synergistic potential of AI and human intelligence, coining the term "high-performance medicine" to describe the future of personalized, precise care. Moreover, Beam and Kohane[22] further highlighted the importance of integrating big data and machine learning tools into healthcare workflows, noting that the scalability of these technologies can improve both outcomes and efficiency.

The rapid pace of innovation is evident, with AI applications in diagnostics, drug development, and remote patient monitoring advancing swiftly, as noted by Heaven.[23] These advances are particularly crucial for enhancing the delivery of healthcare, as outlined by Reddy *et al.*,[24] who discussed both the opportunities and limitations of AI in real-world systems. Similarly, Yu *et al.*[25] provided a more comprehensive review of the challenges in implementing AI technologies, including regulatory hurdles, ethical concerns, and data privacy considerations.

Deep learning, a subset of AI, has become particularly important in medical applications. Esteva *et al.*[26] proposed a practical guide to deep learning tools in healthcare, demonstrating how models, such as convolutional neural networks are now used to interpret medical images and analyze unstructured clinical data. In low-resource settings, AI also holds great promise. Wahl *et al.*,[27] explored how it can be used to reduce health disparities and enhance access to care in underrepresented regions.

The COVID-19 pandemic has accelerated digital adoption in healthcare, including the deployment of AI for surveillance, diagnostics, and modeling. Keesara *et al.*,[28] described this shift as a digital revolution catalyzed by the pandemic. AI's clinical utility is already evident in specialties, such as cardiology, where Dilsizian and Siegel[29] demonstrated its efficacy in cardiac imaging diagnostics.

However, ethical concerns remain. Obermeyer and Emanuel[30] critically examined racial bias in healthcare algorithms, showing how even data-driven tools can perpetuate inequities. This is echoed by Holmes *et al.*,[31] who advocated for a more equity-focused approach to AI deployment, especially in global health contexts.

The role of big data is central to these transformations. Chen and Chen[32] explored how AI-powered analytics can support public health interventions, while Miotto *et al.*[33] proposed "deep patient," an unsupervised model that can accurately predict health outcomes from electronic health records. The global health crisis sparked by COVID-19 further underlined the importance of rapid data modeling and response strategies, as reviewed by Wang *et al.*[34] In addition, LeCun *et al.*[35] laid the theoretical foundation of deep learning technologies, offering insights into neural networks that now underpin most AI systems in healthcare. Building on this, Wang *et al.*[36] emphasized the organizational advantages of big data analytics in hospital and system-level planning.

Furthermore, the use of AI in specific clinical applications is highlighted by Krittanawong *et al.*,[37] who detailed how deep learning enhances cardiovascular risk prediction and diagnostics. From a practical perspective, Davenport and Kalakota[38] and Hinton[39] underscored the growing integration of AI into clinical routines, making a case for workflow optimization and clinical support tools. Finally, the future of AI in healthcare depends on its alignment with human oversight and decision-making. Shortliffe and Sepúlveda[40] highlighted the importance of clinical decision support systems that enhance, rather than replace, physician judgment, ensuring that ethical and contextual considerations remain central to care.

Together, these works establish a comprehensive framework for understanding the potential, progress, and pitfalls of AI in global healthcare systems. The literature collectively supports the premise that, when implemented thoughtfully and equitably, AI holds transformative power to reshape medicine, improve patient outcomes, and reduce disparities worldwide.

This study proposed a novel, AI-driven forecasting framework that combines generative AI (e.g., generative adversarial networks [GANs]) with conventional machine learning methods, such as support vector machines (SVMs) and autoregressive integrated moving average (ARIMA) models. By leveraging generative AI, we aimed to augment incomplete or imbalanced datasets—particularly those from underrepresented regions—thereby enabling more inclusive and equitable global health expenditure forecasting. The proposed methodology not only improves prediction accuracy but also uncovers latent patterns and clusters of countries based on spending behavior. This hybrid approach enhances the interpretability and relevance of predictions for LMICs, empowering policymakers with reliable, data-driven insights to allocate resources more effectively and equitably.

2. Methods

The data utilized in this study were retrieved from the database of the World Bank Group (https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?most_recent_year_desc=true&locations=1W), which retrieved its data from the World Health Organization's Global

Health Expenditure Database (GHED). The GHED is an integrated database comprising a total of 192 countries. A subset of 168 countries was selected in this study. The data were represented in a wide format, comprising country-level healthcare expenditures as a percentage of GDP, collected over the years from 2000 to 2021. In the selected data, the highest health expenditure as a percentage of GDP occurred in Nauru in 2007 (24.23%), whereas the lowest ever was recorded in Qatar in 2011 (1.60%). The US consistently allocated the highest share of GDP spent on healthcare throughout the study period, with a mean of 15.66%, except in 2001, 2007, and 2008 (when Nauru exceeded the US), and in 2014 and 2015 (when Sierra Leone exceeded the US). In contrast, despite maintaining high standards of medical care, Qatar's expenditure on healthcare is one of the lowest among all other countries in the world, averaging 2.61%. This dataset was subsequently used for machine learning classification.

Data from 24 countries were selected from the above dataset and transposed. The new dataset comprised primarily countries from North America and Europe, along with a few Asian countries, including India, China, and Iran. The resulting data are a time series indicating the health expenditure of 24 countries from 2000 to 2021. This dataset was used to predict the health expenditure of selected countries in 2025.

The healthcare expenditure data used in this study were expressed as a percentage of GDP, which inherently normalizes for inflation at the national level by relating spending to the overall economic output. As such, an explicit inflation adjustment was not required for this analysis. Nonetheless, we acknowledge that inflation can still indirectly influence healthcare costs, and future work should consider integrating inflation-adjusted absolute spending values. The following methods were employed to analyze and predict future healthcare expenditures.

2.1. Hierarchical clustering

A hierarchical clustering approach was employed in this study. The healthcare expenditure dataset was organized by nations nested within broader geographic or economic regions, using the city block distance metric and average linkage methods. This structure allowed the understanding of both country-specific variations and regional or global trends over time. In this hierarchical framework, the assumptions were:

(i) Level 1 represents the annual healthcare expenditure for each country, varying by year
(ii) Level 2 captures regional-level or income-based groupings.

The model is represented by Equation I:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij} \tag{I}$$

Where Y_{ij} represents the healthcare expenditure for country j in year i, X_{ij} denotes time-specific covariates, and ε_{ij} is the error term. The intercept β_{0j} and slope β_{1j} terms may vary across countries, allowing the capturing of both the overall trend and country-specific deviations.

2.2. SVMs

The high dimension of the data remained a consistent challenge in employing different algorithms. To reduce the complexity, the SVM model was used for predictive analysis due to its strong performance in classification and regression tasks. SVM aims to identify an optimal hyperplane in higher dimensions that maximally separates data points from different classes or accurately fits the data in regression tasks. For classification, the SVM model constructs a decision boundary that maximizes the margin between data points from distinct classes, thereby minimizing classification error and enhancing generalization to new data.

2.3. ARIMA

In this study, the ARIMA model was employed to analyze and forecast time-dependent trends in healthcare expenditures across countries. The ARIMA model is a widely used statistical method for analyzing time series data, especially when a series exhibits temporal autocorrelation. The model is particularly advantageous for handling non-stationary data by integrating differencing, autoregressive, and moving average components, making it suitable for our dataset, which spans multiple years. Based on the autocorrelation function and partial autocorrelation function diagnostics, initial ARIMA model configurations were identified and fitted to the data. Model performance was assessed using standard evaluation metrics, including the Akaike information criterion (AIC) and Bayesian information criterion (BIC), to select the most parsimonious model with adequate predictive accuracy.

2.4. Generative AI

Generative AI, a subset of AI that focuses on generating new data instances resembling existing data, offers innovative solutions for complex problems in health economics. Unlike traditional predictive models, generative AI models, such as GANs and variational autoencoders (VAEs), excel in simulating realistic scenarios and exploring data-driven solutions, making them a valuable tool for understanding and forecasting healthcare expenditure under various economic, demographic, or policy-driven conditions. By leveraging models, such as GANs or conditional VAEs, researchers can generate plausible future trends based on historical data and specific input variables. For example, these models can simulate how an economic downturn, a new healthcare policy, or a pandemic may affect spending patterns across regions or income groups. Such simulations enable policymakers to explore potential outcomes before implementing changes, enabling proactive planning and risk mitigation.

3. Results

3.1. Machine learning

3.1.1. Hierarchical clustering

The results of the hierarchical clustering showed six clusters with similar healthcare expenditure from 2003 to 2021 (Figure 1). The information of each cluster, including size and countries, is depicted as follows:

(i) Cluster 0 (30 countries): The UK, Armenia, Canada, Austria, Switzerland, Netherlands, Sweden, Portugal, Belgium, Denmark, Japan, Spain, Malta, Australia, Finland, Lesotho, Norway, New Zealand, Maldives, Serbia, Brazil, Iceland, El Salvador, Argentina, Bosnia and Herzegovina, Slovenia, Namibia, Italy, Uruguay, and Greece.

(ii) Cluster 1 (46 countries or regions): Timor-Leste, Lebanon, Nicaragua, Panama, Czechia, Cyprus, Chile, Republic of Korea, Honduras, Mozambique, Latvia, Colombia, Bulgaria, North Macedonia, Georgia, Andorra, Ecuador, South Africa, Guinea-Bissau, Latin America and Caribbean (excluding those high-income countries), Bolivia, Croatia, Barbados, Paraguay, Tajikistan, Ukraine, San Marino, Israel, Lithuania, Slovak Republic, Costa Rica, Cambodia, Estonia, Malawi, Hungary, Rwanda, Jordan, Albania, Eswatini, Tunisia, Guatemala, Belarus, Poland, Botswana, Mexico, and Iran.

(iii) Cluster 2 (51 countries or regions): Uzbekistan, Russia, Jamaica, The Bahamas, Trinidad and Tobago, Mongolia, Cabo Verde, Samoa, Zambia, Dominica, Romania, Mauritius, Burkina Faso, Comoros, Caribbean small states, Tonga, Peru, Saudi Arabia, Philippines, Niger, Morocco, Grenada, Suriname, Luxembourg, Turkmenistan, Togo, Algeria, Nepal, China, Seychelles, Chad, Belize, Guyana, Dominican Republic, Uganda, Egypt, Vietnam, Türkiye, Kenya, Mali, Senegal, Bahrain, Ghana, Eritrea, Monaco, Madagascar, Haiti, Tanzania, Ethiopia, The Gambia, and Sudan.

(iv) Cluster 3 (1 country): The US.

(v) Cluster 4 (4 countries): Palau, Cuba, Germany, and France.

Cluster 5 (27 countries): Kuwait, Myanmar, Singapore, Fiji, United Arab Emirates, Iraq, Thailand, Azerbaijan, Malaysia, Vanuatu, Oman, Mauritania, Nigeria, Sri Lanka, Kazakhstan, Bhutan, Cameroon, Guinea, Indonesia, India, Angola, Pakistan, Qatar, Djibouti, Gabon, Benin, and Bangladesh.

It is noticeable that Clusters 0, 3, and 4 comprised developed economies, including the US, Japan, the UK, Nordic countries, Germany, and France, indicating that developed countries allocate a higher percentage of GDP than developing countries and emerging economies. Other

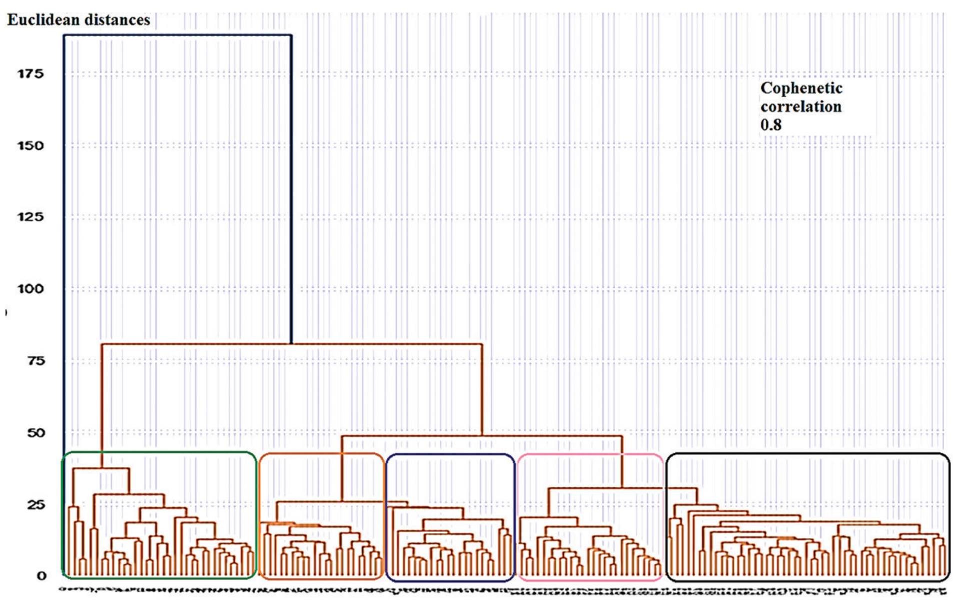

Figure 1. Hierarchical clustering graph of 158 countries
Notes: Green: Cluster 0; Orange: Cluster 1; Blue: Cluster 2; Pink: Cluster 3; Black: Cluster 4.

clusters consisted of countries from Eastern Europe, Latin America, Africa, the Caribbean, Asia, and the Middle East, implying that most countries in these regions tend to allocate a lower percentage of GDP to healthcare than other developed countries.

3.1.2. SVMs

The linear SVM, polynomial SVM, and Gaussian radial basis function (RBF) SVM models demonstrated accuracies of approximately 77, 71, and 81%, respectively. The graphs of 158 countries are shown in Figure 2. The US is in one cluster with a substantial difference from other countries. The linear SVM results showed straight lines across classes, indicating clear linear separations. However, it might also imply an oversimplification of the relationship across classes.

The polynomial SVM results demonstrated slightly curved boundaries compared to linear boundaries, suggesting the added complexity of the polynomial transformation was unnecessary. Moreover, the Gaussian model generated more complex and non-linear boundaries with curved and localized decision regions, implying potentially better capacity in capturing the actual association across classes. Its superior accuracy to other methods suggests that the data contain patterns, such as complicated overlapping classes and non-linearly separable data.

3.2. Predictive modeling

3.2.1. ARIMA

To analyze health expenditure trends across countries, a simple ARIMA model was initially fitted to each country. The model was specified with an order of (1, 1, 1), where the parameters represent the following: One lag for the autoregressive term, first-degree differencing to address non-stationarity, and one lag for the moving average term. This configuration utilized the values from one step prior for both autoregressive and moving average components while assuming stationarity of the time series as the primary requirement for ARIMA models.

However, visual inspection indicated that the data exhibited non-stationary behavior. To address this, first-degree differencing (d = 1) was applied, ensuring the data met the stationarity assumption. Figure 3 visualizes the predicted health expenditure of 25 selected countries in a map.

3.2.2. Predicted values

Table 1 presents the predicted health expenditures for the year 2025 across all selected countries, using both the

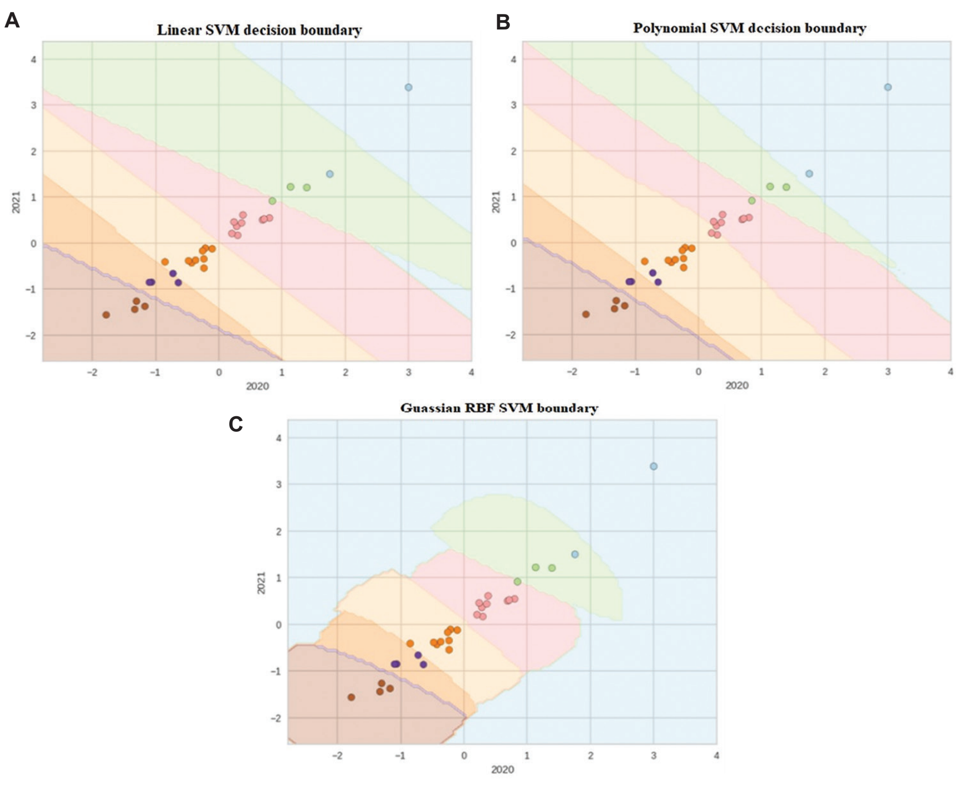

Figure 2. Decision boundaries of (A) linear, (B) polynomial, and (C) Gaussian radial basis function support vector machine models for 2020–2021

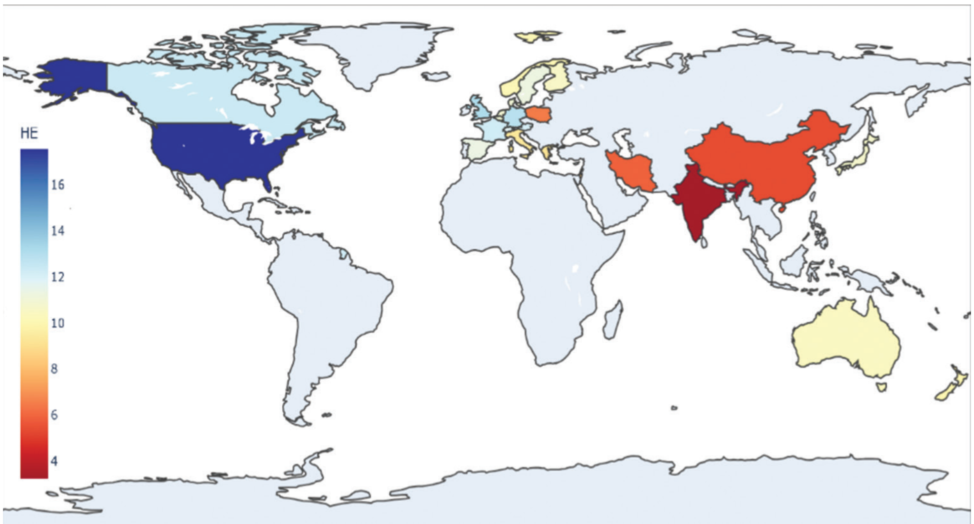

Figure 3. Predicted health expenditure of selected countries for 2025 using a simple autoregressive integrated moving average model (p = 1, d = 1, q = 1)

Table 1. Predicted health expenditure as percentages of GDP for 2025 using a simple ARIMA model (p=1, d=1, q=1) and a multi-model ARIMA

Country	Simple ARIMA (%)	Multi-model ARIMA (%)
United States	17.57	17.53
United Kingdom	13.13	12.36
Switzerland	11.77	11.80
Sweden	11.27	11.25
Spain	11.33	10.74
Poland	6.46	6.24
Norway	10.15	10.08
New Zealand	10.02	9.83
Netherlands	10.95	11.29
Japan	10.78	10.82
Italy	9.36	9.09
Israel	8.06	7.90
Iran	5.88	5.77
India	3.25	3.28
Greece	9.16	9.17
Germany	12.89	12.93
France	12.25	12.31
Finland	10.59	10.49
Denmark	11.03	10.82
China	5.36	5.38
Canada	12.38	12.42
Belgium	10.99	11.04
Austria	12.12	12.10
Australia	10.46	10.54

Abbreviations: ARIMA: Autoregressive integrated moving average; GDP: Gross domestic product.

simple ARIMA model and the enhanced multi-model approach. While the simple ARIMA model demonstrated favorable AIC scores, the multi-model approach achieved further reductions in AIC by fitting ARIMA models with all possible parameter combinations. This optimization led to losing their autoregressive component for most countries, except for Canada, New Zealand, Italy, and Poland.

The differences in predicted expenditures between the two approaches ranged from as little as 0.01 (for Greece) to 0.76 (for the UK). These results highlight the utility of the multi-model approach in refining parameter selection to enhance predictive accuracy. Subsequently, the fitted models were used to predict health expenditure for the next 4 years, with the results summarized in Figure 4, which presents the comprehensive prediction graphs for each country.

3.3. Generative AI

GANs play a transformative role in augmenting sparse and underrepresented datasets by generating synthetic yet realistic data. In the context of healthcare expenditure prediction, GANs can create additional data points for countries or regions where reliable data are scarce.

Figure 5 illustrates the progression of GAN training. In Figure 5A, epoch 0 represents the state before training, where the generated data (red points) were scattered randomly, showing no resemblance to the real data (blue points). This stage highlights the GAN's initial random generation, as the generator has not yet learned the underlying patterns of the real dataset.

The latent space dimension and batch size of trained data were 150 and 64, respectively. For epoch 0, the

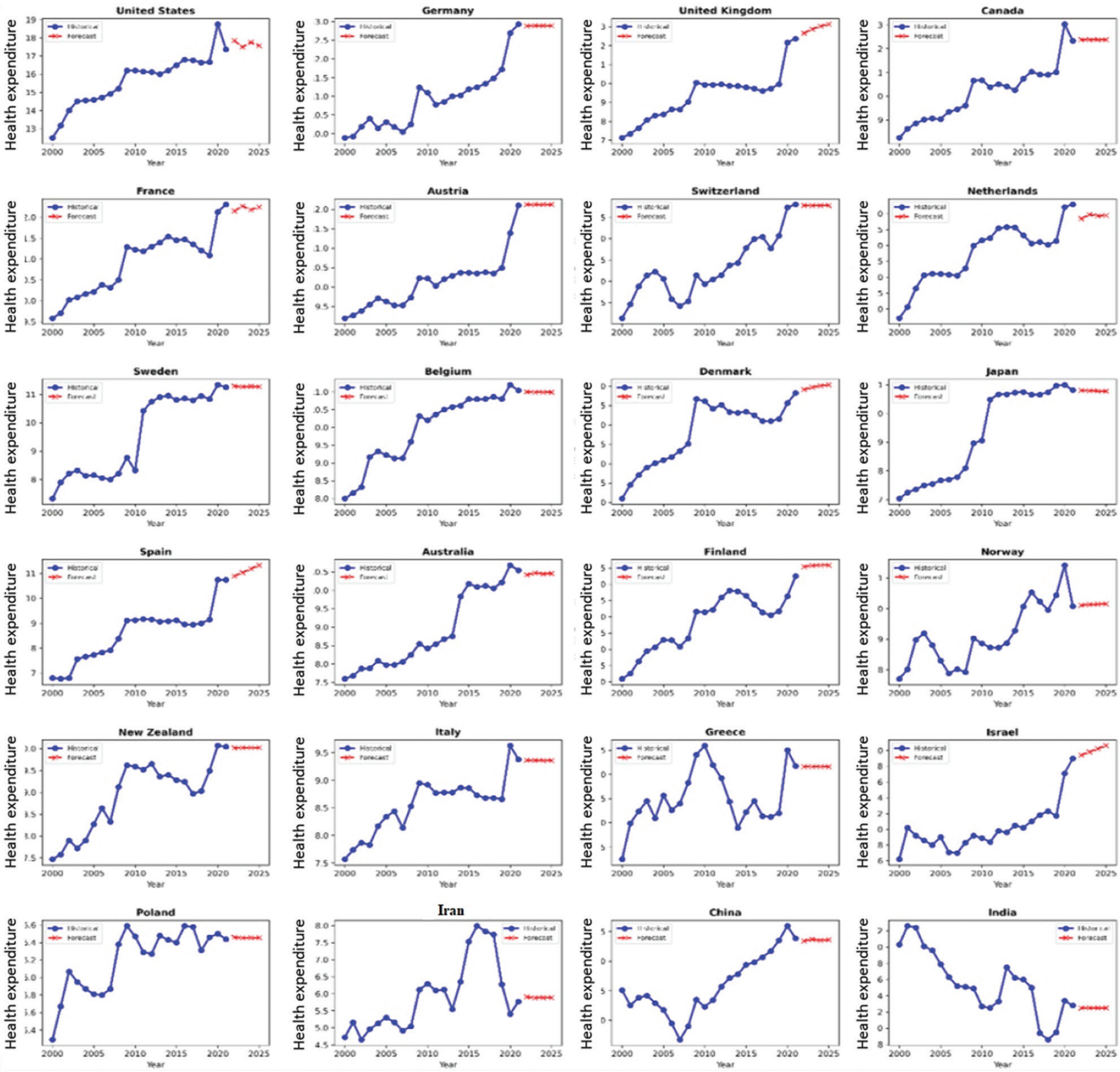

Figure 4. Predicted health expenditure of selected countries for 2022–2025 with a simple autoregressive integrated moving average model (p = 1, d = 1, q = 1)

discriminator loss was 1.396, implying that the generator makes many errors in classifying real and simulated samples. Meanwhile, the generator loss was 0.783, indicating the generation of fake data. In Figure 5B, epoch 2,000 presented the GAN's capability after extensive training. Here, the generated data aligned closely with the real data, demonstrating the generator's success in learning and replicating the distribution of healthcare expenditure patterns. This alignment indicates the potential of GANs to generate high-quality synthetic data, which can be used to address data sparsity issues in underrepresented regions or time periods. By filling gaps in datasets, GANs enable more accurate predictive modeling and equitable analysis of healthcare expenditures globally. Increasing the number of epochs to 2,000 reduced the discriminator and generator losses to 0.67 and 0.73, respectively.

Generative AI brings transformative advantages to healthcare expenditure analysis by addressing the limitations of traditional predictive models. One significant benefit is its ability to handle incomplete or imbalanced datasets through data augmentation. By generating synthetic yet realistic data, generative AI fills gaps in underrepresented

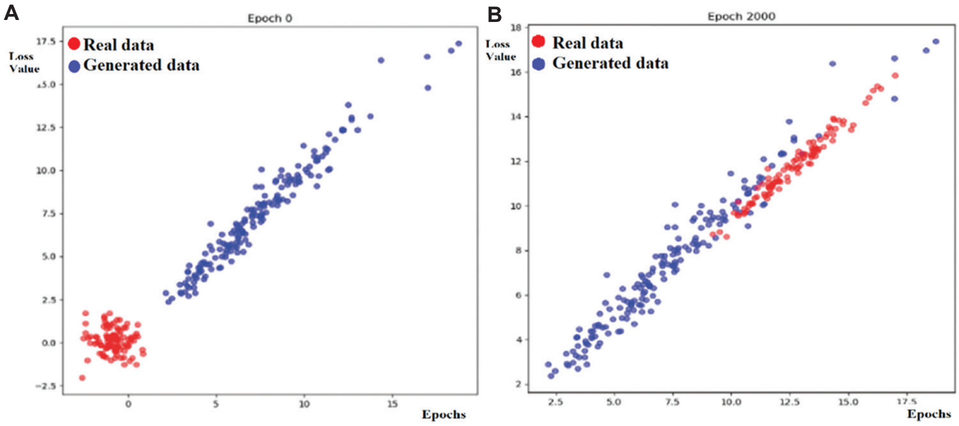

Figure 5. Outputs of generative adversarial networks at (A) epoch 0 and (B) epoch 2,000

regions, ensuring that models trained on such data are more inclusive and generalizable. This capability enhances the reliability of global healthcare expenditure predictions, particularly for low-income countries with limited historical data, thus fostering equitable decision-making.

4. Discussion

Despite its transformative potential, implementing generative AI in healthcare expenditure analysis faces significant challenges, particularly related to data quality and privacy. Generative AI models rely heavily on large, high-quality datasets to produce accurate and reliable outputs. However, healthcare expenditure data often suffer from inconsistencies, incomplete records, and regional disparities in reporting. For underrepresented countries or regions with scarce data, the generative AI model's ability to produce meaningful synthetic data might be compromised. In addition, concerns around data privacy and compliance with regulations, such as the General Data Protection Regulation and Health Insurance Portability and Accountability Act, pose obstacles to obtaining and utilizing sensitive healthcare-related information for training models.

Another major challenge is the computational complexity and resource requirements of generative AI models. Training sophisticated models, such as GANs or VAEs, requires substantial computational power, specialized infrastructure, and expertise, all of which may not be readily available in low-resource settings. Moreover, generative AI models often lack interpretability, making it difficult for policymakers and stakeholders to fully trust or understand the outputs. The potential for generating unrealistic or biased synthetic data further emphasizes the need for rigorous validation and oversight mechanisms. Overcoming these challenges requires collaboration across AI researchers, policymakers, and healthcare experts to ensure ethical, efficient, and equitable implementation of generative AI solutions. Validation processes include:

Table 2. Validation methods and the accuracy of all models

Model	Data split	Validation method	Results
SVM (linear)	70% train 15% validation 15% test	Cross-validation (k-fold)	Approximately 77% accuracy
SVM (polynomial)	70% train 15% validation 15% test	Cross-validation (k-fold)	Approximately 71% accuracy
SVM (RBF)	70% train 15% validation 15% test	Cross-validation (k-fold)	Approximately 81% accuracy (highest among SVM models)
ARIMA (simple)	Time-series split (train/test by year)	AIC, BIC, holdout year prediction	Accurate short-term forecast: AIC/BIC optimized
ARIMA (multi-model)	Time-series split (train/test by year)	Grid search for (p, d, q) combinations	Improved model fit for most countries
GAN	- (unsupervised)	Visual convergence and loss monitoring	Generator loss reduces to 0.73 and discriminator loss reduces to 0.67 by epoch 2000

Abbreviations: AIC: Akaike information criterion; ARIMA: Autoregressive integrated moving average; BIC: Bayesian information criterion; GAN: Generative adversarial network; RBF: radial basis function; SVM: Support vector machine.

(i) ARIMA models: AIC and BIC were used to select optimal parameters and evaluate forecast accuracy by comparing predicted values versus actual expenditures in a holdout year.
(ii) SVM models: Cross-validation was applied, and classification accuracy was reported across different kernels. The RBF kernel achieved the highest accuracy at 81%.
(iii) GAN models: Model fidelity was evaluated using

visual comparison of generated versus actual data distributions, and loss curves were monitored, with both generator and discriminator losses decreasing over training.

The validation and accuracy of the training, test, and validation sets are summarized in Table 2. A standard approach was used for data partitioning. The dataset was split into training (70%), validation (15%), and test (15%) sets based on a stratified sampling method to maintain the distribution of healthcare expenditure across different country groups. For the SVM models, cross-validation was applied during training to optimize model parameters and mitigate overfitting.

5. Conclusion

This study demonstrates the potential of integrating generative and traditional AI techniques to forecast healthcare expenditures globally, with a strong emphasis on addressing data sparsity and equity in LMICs. By applying a hybrid framework of GANs for synthetic data augmentation, SVMs for classification, and ARIMA for temporal forecasting, the robustness and accuracy of predictions were enhanced, particularly in data-scarce settings. The findings highlight that generative AI can mitigate the limitations of incomplete datasets, enabling a more representative modeling of global expenditure patterns. This methodological advancement is especially impactful for LMICs, where historical health data may be inconsistent or unavailable. By simulating realistic expenditure trends, the proposed framework equips policymakers with actionable forecasts tailored to their unique economic and demographic conditions. Ultimately, this work contributes to the development of a scalable and replicable AI-based model that supports equitable, evidence-based decision-making in global health finance, bridging a critical gap between advanced predictive technologies and real-world health policy needs.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Taegeon Yu, Wen-Shan Liu, Payam Norouzzadeh, Eli Snir, Bahareh Rahmani
Investigation: Daipayan Bera, Abbas Maazallahi, Roschlynn Dsouza, Francina Pali, Eli Snir, Bahareh Rahmani
Methodology: Taegeon Yu, Daipayan Bera, Abbas Maazallahi, Roschlynn Dsouza, Francina Pali, Wen-Shan Liu, Bahareh Rahmani
Writing – original draft: Taegeon Yu, Daipayan Bera, Abbas Maazallahi, Roschlynn Dsouza, Francina Pali, Bahareh Rahmani
Writing – review & editing: Payam Norouzzadeh, Eli Snir, Bahareh Rahmani

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

You may download the data from https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?most_recent_year_desc=true&locations=1W.

References

1. *Global Medical Trends Survey*; 2024. Available from: https://www.wtwco.com/en-cz/insights/2023/11/2024-global-medical-trends-survey [Last accessed on 2025 Sep 18].
2. Meijer CD, Wouterse B, Polder J, Koopmanschap M. The effect of population aging on health expenditure growth: A critical review. *Eur J Ageing*. 2023;10(4):353-361.

   doi: 10.1007/s10433-013-0280-x
3. Global Burden of Disease Health Financing Collaborator. Past, present, and future of global health financing: A review of development assistance, government, out-of-pocket, and other private spending on health for 195 countries of 1995–2050. *Lancet*. 2019;393(10187):2233-2260.

   doi: 10.1016/S0140-6736(19)30841-4
4. Young AJ, Terry RF, Rottingen JA, Viergever RF. Global trends in health research and development expenditures-the challenge of making reliable estimates for international comparison. *Health Res Policy Syst*. 2015;13(1):7.

   doi: 10.1186/1478-4505-13-7
5. Wilensky G, Gordon H, Sun MD, *et al*. Global trends in biomedical R&D expenditures. *N Engl J Med*. 2014;370(1):3-6.

   doi: 10.1056/NEJMp1311068
6. Wager E, Rakshit S, Cox C. *What Drives Health Spending in the U.S. Compared to other Countries?* Washington, D.C: KFF Health Cost; 2024.
7. McGough M, Winger A, Kurani N, Cox C. *How Much is Health Spending Expected to Grow?* Washington, D.C: KFF Health Cost; 2024.

8. Glassman A, Keller JM, Smitham E. *The Future of Global Health Spending Amidst Multiple Crises*. United States: Center for Global Development; 2023.

9. Baltagi BH, Lagravinese R, Moscone F, Tosetti E. Health care expenditure and income: A global perspective. *Health Econ*. 2017;26(7):863-874.

doi: 10.1002/hec.3424

10. Gascon AE, Mico-Sanz JL, Ripolles AC. Global evidence of environmental and lifestyle effects on medical expenditures across 154 countries. *Prev Med Rep.* 2022;30:102036.

doi: 10.1016/j.pmedr.2022.102036

11. Chen F, Chen Z. Cost of economic growth: air pollution and health expenditure. *Sci Total Environ*. 2021;755(Pt 1):142543.

doi: 10.1016/j.scitotenv.2020.142543

12. Marley R. Top 10 current trends expected to transform healthcare in 2024. Healthcare Transformers. 2024. Available from: https://healthcaretransformers.com/digital-health/current-trends/current-trends-expected-to-transform-healthcare-in-2024/ [Last accessed on 2025 Sep 18].

13. Bertl M, Ross P, Draheim D. Systematic AI support for decision-making in the healthcare sector: obstacles and success factors. *Health Policy Technol.* 2023;12(3):100748.

14. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value Health.* 2019;22(7):808-815.

doi: 10.1016/j.jval.2019.02.012

15. Langenberger B, Schulte T, Groene O. The application of machine learning to predict high-cost patients: A performance-comparison of different models using healthcare claims data. *PLoS One*. 2023;18(1):e0279540.

doi: 10.1371/journal.pone.0279540

16. Obermeyer Z, Emanuel EJ. Predicting the future, big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375(13):1216-1219.

doi: 10.1056/NEJMp1606181

17. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: A practical introduction. *BMC Med Res Methodol*. 2019;19(1):64.

doi: 10.1186/s12874-019-0681-4

18. Khan JR, Awan N, Awan N, Islam M, Muurlink O. Healthcare capacity, health expenditure, and civil society as predictors of COVID-19 case fatalities: A global analysis. *Front Public Health*. 2020;8:347.

doi: 10.3389/fpubh.2020.00347

19. Mathauer I, Oranje M. Global medical trends survey report. *Bull World Health Organ*. 2024;102(3):216-224.

20. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-1358.

doi: 10.1056/NEJMra1814259

21. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. *Nat Med*. 2019;25:44-56.

doi: 10.1038/s41591-018-0300-7

22. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317-1318.

doi: 10.1001/jama.2017.18391

23. Heaven WD. AI is changing medicine fast. *Nature*. 2023;614:234-237.

24. Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. *J R Soc Med.* 2019;112(1):22-28.

doi: 10.1177/0141076818815510

25. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2:719-731.

doi: 10.1038/s41551-018-0305-z

26. Esteva A, Robicquet A, Ramsundar B, *et al*. A guide to deep learning in healthcare. *Nat Med.* 2019;25:24-29.

doi: 10.1038/s41591-018-0316-z

27. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Health*. 2018;3:e000798.

doi: 10.1136/bmjgh-2018-000798

28. Keesara S, Jonas A, Schulman K. Covid-19 and health care's digital revolution. *N Engl J Med*. 2020;382:e82.

doi: 10.1056/NEJMp2005835

29. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: Harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep.* 2014;16(1):441.

doi: 10.1007/s11886-013-0441-8

30. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453.

doi: 10.1126/science.aax2342

31. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *Lancet*. 2020;395(10236):1579-1586.

doi: 10.1016/S0140-6736(20)30226-9

32. Chen M, Hao Y, Cai Y, Wang Y, Sun X. The role of big data and artificial intelligence in public health. *Biomed Res Int.* 2020;2020:1-10.

33. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: An unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep.* 2016;6:26094.

34. Wang L, Wang Y, Ye D, Liu Q. Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *Int J Antimicrob Agents*. 2020;55(6):105948.

doi: 10.1016/j.ijantimicag.2020.105948

35. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.

    doi: 10.1038/nature14539

36. Wang Y, Kung L, Byrd TA. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. *Technol Forecast Soc Change*. 2018;126:3-13.

37. Krittanawong C, Johnson KW, Rosenson RS, *et al*. Deep learning for cardiovascular medicine: A practical primer. *Eur Heart J*. 2019;40(25):2058-2073.

    doi: 10.1093/eurheartj/ehz056

38. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-98.

    doi: 10.7861/futurehosp.6-2-94

39. Hinton G. Deep learning technology with the potential to transform health care. *JAMA*. 2018;320(11):1101-1102.

    doi: 10.1001/jama.2018.11100

40. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA*. 2018;320(21):2199-2200.

    doi: 10.1001/jama.2018.17163