frontiers | Research Topics

Health economics, medical technology and artificial intelligence

Edited by
Yingying Xu and Ke Yan

Coordinated by
Zengshou Dong

Published in
Frontiers in Public Health
Frontiers in Oncology

ISSN 1664-8714
ISBN 978-2-8325-4535-5
DOI 10.3389/978-2-8325-4535-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Health economics, medical technology and artificial intelligence

Topic editors
Yingying Xu — Beihang University, China
Ke Yan — National University of Singapore, Singapore

Topic coordinator
Zengshou Dong — Taiyuan University of Science and Technology, China

Citation
Xu, Y., Yan, K., Dong, Z., eds. (2024). *Health economics, medical technology and artificial intelligence*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-4535-5

Table of contents

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 20 January 2023
DOI 10.3389/fpubh.2023.1121154

OPEN ACCESS

EDITED BY
Yingying Xu,
Beihang University, China

REVIEWED BY
Yifei Cai,
Teesside University, United Kingdom
Giray Gozgor,
Istanbul Medeniyet University, Türkiye
Umer Niaz,
Punjab Group of Colleges, Pakistan

*CORRESPONDENCE
Jinyan Hu
✉ hwx@sdu.edu.cn

SPECIALTY SECTION
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

RECEIVED 11 December 2022
ACCEPTED 04 January 2023
PUBLISHED 20 January 2023

CITATION
Rong Y and Hu J (2023) How can green credit decrease social health costs? The mediating effect of the environment.
Front. Public Health 11:1121154.
doi: 10.3389/fpubh.2023.1121154

How can green credit decrease social health costs? The mediating effect of the environment

Yanbo Rong and Jinyan Hu*

School of Economics, Qingdao University, Qingdao, Shandong, China

Green credit plays an important role in environmental protection and residents' health. This paper discusses the impact path of green credit on social health costs with the help of a quantile regression. The implementation of a green credit policy can decrease social health costs in China, and green credit works best in the economically developed Eastern region. As the quantile increases, so gradually does the absolute value of the green credit coefficient. This result proves that for provinces with rich per capita financial health expenditures, green credit plays a greater role in decreasing social costs, a conclusion also supported by our robustness test. In addition, we find that environmental pollution plays a mediating role in the path of green credit affecting health, and this finding is verified in the green credit and health general equilibrium model. Based on these findings, the government should encourage the active innovation of green credit products, and the banking industry should develop personalized green credit products for specific pollutant types or industries while decreasing government pressure.

KEYWORDS

green credit, social health costs, environmental pollution, quantile regression (QR), mediating role

1. Introduction

Underlying rapid economic development is the massive cost of environmental resources and medical and health care. According to the World Bank, although energy consumption per unit of GDP has decreased, medical expenditures have increased rapidly. At present, pollution is the world's largest environmental risk factor for disease and premature death, especially in low- and middle-income countries. In 2019, 9 million people died from pollution, equivalent to 1/6 of the global death toll[1]. Pollution and pollution-related consequences are the main inducements that have a massive impact on health (1), having a negative effect on the mental health of the public, and decreasing subjective wellbeing and mental health. Environmental pollution is closely related to the allocation of credit resources. Green credit supports low energy consumption, low emissions, low pollution, and high efficiency business behaviors, the last of which is an important practice of guiding green development through the rational allocation of credit funds (2). Improving the allocation of credit capital to green industries can curb the impact of environmental pollution on residents' health, which is conducive to achieving win–win economic and environmental goals (3, 4).

China's environmental pollution has always been very serious. Outdoor air pollution has become the fourth leading lethal risk factor (5), and the number of deaths due to environmental pollution in China ranks second globally (6). As China has industrialized, infrastructure construction and the heavy chemical industry have consumed large amounts of mineral resources and fossil fuels and have had an important impact on people's health. Thus, the health of Chinese households is directly related to the rapid development of resource-intensive

1 Data source: The Lancet.

industries as a result of traditional credit (7). China proposed supply-side structural reform in 2015, hoping to optimize the allocation of financial resources and improve residents' health. The 13th 5-Year Plan clearly included structural adjustment and environmental protection as the overall goal of macro control. The 19th National Congress of the CPC made it clear that "we should treat the ecological environment like life" and listed environmental protection as one of the "three tough battles." Under the current multiple difficulties of prominent financial and credit mismatch, worsening ecological environment and increasing household health expenditures, how to improve the quality of economic growth and achieve sustainable growth are core issues for China's future economic development.

This paper makes three main contributions. Firstly, we consider the new path of environmental pollution to explain the impact of green credit on social health costs. The literature has generally been based on the impact of green credit on enterprises' green technology innovation to improve environmental quality. Alternatively, some research has been done on the harm environmental pollution does to residents' health. The promotion of green credit in financial fields can decrease financial pressure and release further funds for health. More importantly, residents' health can be improved by improving environmental quality. This fact also conforms to green credit and health general equilibrium theory, which verifies the intermediary role played by the environment. Secondly, this paper uses a quantile regression (QR) to test the impact of green credit on social health costs under different quantiles. The independent variable has a specific regression curve on each characteristic quantile of the dependent variable, which is the advantage of this model. For provinces with higher per capita financial health expenditures, green credit plays a greater role in decreasing social costs. Finally, our study fully considers heterogeneity. The impact of green credit on social health costs shows an obvious geographical ladder distribution, with the largest effect in the Eastern region, followed by the Central region, and an as-yet-undiscovered effect in the Western region. This conclusion is helpful for banks as they issue targeted green credit policies and decreases social health costs in underdeveloped areas.

The rest of this study is organized as follows. Section 2 covers the literature review. Section 3 outlines the theoretical model. Section 4 describes the methodology. Section 5 describes the data. Section 6 presents a discussion of the empirical findings. Section 7 summarizes the results and discusses some policy implications.

2. Literature review

Green credit can affect national health spending in two ways. On the one hand, green credit can optimize the industrial structure and cause more people to engage in tertiary industries, thus directly improving residents' health levels. On the other hand, green credit can improve environmental conditions and play a positive role in public health. Therefore, considering the intermediary effect of the environment, this paper expands the existing literature in three ways: green credit and health, green credit and the environment, the environment and health.

2.1. Green credit and health

As an important guarantee for the real economy, sustainable development, and a country's core competitiveness, credit business can play an indispensable role in building a community of human health. Frederik et al. (8) found that the debt burden is steadily increasing of low- and middle-income countries, coupled with the economic recession of COVID-19, which combined might nullify the necessary asset health expenditures. Tuohy et al. (9) showed that the impact of private financing on publicly funded health care systems depends on the construction of the relationship between public and private financing. The authors think that the increase in private health expenditures partly replaces public finance. Leatherman (10) showed that microfinance institutions have advantages in developing health financing programs, which can expand poor people's existing choices, and prevent the risk of poverty caused by disease. Maurya and Asher (11) revealed that India spent 3.7% of its GDP on health care, but the health care results were not commensurate with the expenditures. Therefore, increasing public health expenditures alone does not improve health outcomes unless the inefficiency of existing public and private health financing arrangements is addressed.

The research in China has drawn a relatively consistent conclusion: green credit has a positive impact on health. Liu and Guo (12) discussed the significant positive effect of inclusive finance on public health based on China's data. Hu et al. (13) found that China's green credit has significant effects on the transformation of the industrial structure. Green credit mainly influences population health through the industrial structure. Zhu (14) combined green credit, technological innovation, industrial structure upgrades, and population health; theoretically analyzed the impact of green credit and technological innovation on industrial structure upgrades and population health; and analyzed how green credit affects population health through technological innovation and industrial structure upgrades.

2.2. Green credit and environmental conditions

Green credit has effectively decreased the pollution and energy consumption of high-emission enterprises (15), and the literature has generally concluded that green credit can greatly improve environmental quality. Based on data from the BRICS countries (i.e., Brazil, Russia, India, China, and South Africa), Wang et al. (16) revealed that inclusive finance is a catalyst to promoting the growth of renewable energy investment and decreasing carbon emissions. Zeng et al. (17) found that green finance, EU consumption, and technological innovation performed well in protecting the environment by decreasing carbon emissions. Muhammad et al. (18) used data from South Asia to show that green bonds, decreasing greenhouse gas emissions, and green economic development played an important role in green financial development. Zhang et al. (19) conducted a dynamic relationship study on samples from 49 countries and showed that green finance could effectively mitigate environmental pollution and climate change. Accelerating the development of green finance is the primary driving force for achieving sustainable development.

Researchers have also reached a consensus on green finance in China. Qiao et al. (20) found that China's financial development is positively related to environmental pollution, and further noted that China is in the first stage of an "environmental dividend." To enter the second stage of "sustainable finance," China should increase the reasonable income of "green finance" and establish a unified national carbon trading market. Sun et al. (21) found that China's green credit policy has greatly encouraged enterprises, especially those that rely heavily on external financing, to decrease water pollution. Huang et al. (22) showed a significant positive auto-correlation between green finance and green innovation. Tang et al. (23) discussed the relationship between green finance and the ecological environmental quality of the Yangtze River Economic Belt and found a significant positive impact. Xu and Zhu (24) proved that China's overall green governance index and green financial policies have significantly decreased environmental pollution. In addition, some literature has studied the ways green finance restricts polluting enterprises, which has increased green technology innovations (7, 25). Chen et al. (26) affirmed that financial development can ease the financing constraints faced by innovative activities and promote green technological innovation. Hong et al. (27) considered that green credit guidance mainly restricts green technology innovation by decreasing debt financing rather than through financing constraints. Yang and Zhang (28) thought that implementing green credit policies significantly inhibited the long-term financing of heavily polluting enterprises but allowed heavily polluting enterprises to expand their short-term financing. Chai et al. (29) also used data from Chinese enterprises to support this view.

However, green finance does not invariably improve the environment: at different economic development levels, the influence differs (30). For example, Zhong (31) realized that digital finance's environmental improvement has a threshold, after which an acceleration effect can result.

2.3. Environment and health

Pollution forces humans to face massive health costs and survival threats, leading to the rapid depreciation of health human capital. Thus, pollution constitutes an important source of inequality in economic and social development. The health demand function first proposed by Grossman (32) focused on environmental pollution and residents' health from an economic perspective. Based on this function, Cropper (33) and Wagstaff (34) introduced environmental pollution into the model as an important variable that significantly affects human health. Since then, many studies have affirmed the importance of environmental quality and the role of public health functions in improving public health in cooperation with local governments. Scally and Perkins (35), Welsch (36), and Knibbs et al., (37) found that long-term exposure to high particle matter concentrations increased the risk of children suffering from asthma. Awais and Tariq (38) proved that the building environment was linked to human health through physical activity opportunities, and pollutants such as PM10 also increased risks (39–41). Some literature has discussed the harm environmental pollution poses to residents' physical and psychological wellbeing. Albertini et al. (42) showed that environmental pollution is the major cause, other than age, of the increase in the health depreciation rate and the accelerated decrease in the health stock. Ivanova (43) considered that clean air is a prerequisite for human health and happiness.

The research on China has also drawn a relatively consistent conclusion. Zhao et al. (44) thought that China's massive energy consumption and low energy efficiency have caused severe environmental pollution and posed a great threat to the national health level. The general occurrence of health complications caused by cumulative environmental pollution in China is on the rise. Wang et al. (45) found that external environmental pollution and subjectively perceived pollution are negatively correlated with public health. Specifically, air pollution and domestic waste pollution have significant associations, mainly with public mental health. Wang et al. (46) proved that the environmental health indicators of atmospheric pollution in Central and Eastern China are low, indicating a serious environmental health condition. Chen et al. (47) revealed that the differential heating policy between the north and the south of China was very likely to cause air pollution in the north, which would led to a decrease in residents' average life span in this region by approximately 5.5 years.

Regarding the research on green credit, the existing literature mostly discusses its impact on the green innovation or transformation and upgrading of polluting enterprises, which are activities that improve environmental quality. However, few studies have examined the impact of green credit on health. More studies have been focused on the substitution of financial health expenditures from the private financing perspective, and little research has been conducted from an empirical perspective on the direct and indirect pathways. This paper not only considers the direct impact of green credit on social health costs but it also discusses the mediating effect of environmental pollution, which enriches the existing research.

3. Green credit and health general equilibrium model

According to the family production function (32), a simplified health production function can be expressed as:

$$H = G(X) = G(X_1, X_2, \cdots, X_n) \tag{1}$$

where X_i is the factors influencing health, including environment (E), health expenditures (M0), population structure (P), and other variables (Z). Then, the health production function can be expressed as:

$$H = G(E, M, P) = \Omega \prod E_i^{\alpha_i} \prod M_i^{\beta_i} \prod P_i^{\gamma_i} \prod Z_i^{\lambda_i} \tag{2}$$

where α, β, γ, *and* λ are the corresponding elasticity coefficients, and Ω is the estimated value of initial social health. Taking a logarithm of Equation (2), we obtain:

$$LnH = Ln\Omega + \sum \alpha_i (LnE_i) + \sum \beta_i (LnM_i) + \sum \gamma_i (LnP_i) + \sum \lambda_i (LnZ_i) \tag{3}$$

We expand the financial sector based on neoclassical growth theory with resource constraints. Suppose an economy has two

enterprise types that provide intermediate products. h represents polluting enterprises, which are characterized by heavy assets, such as steel and petrochemical; l represents clean enterprises, which are characterized by light assets and high technological levels, such as electronic components. The production functions of the two enterprise types are:

$$Y_h = N^{\beta_1} K_h^{\beta_2}; \; Y_l = A\,(t)\, K_l^{\gamma} \tag{4}$$

where h enterprises consume natural resources N and material capital K_h, l enterprises do not consume natural resources (actually, they consume fewer, a fact that is simplified here), and use only healthy human capital H and material capital K_l, with technological progress parameter A. In the consideration of technological progress, A is a function of time t. Production is also affected by the environment. Following the assumption of Bowenberg and Simus (48), P represents pollution, and as it deepens, the output decreases more. Thus, the production function is as follows:

$$Y = Y_h^{\alpha_1} Y_l^{\alpha_2} P^{-\alpha_3}, \alpha_1, \alpha_2, \alpha_3 > 0 \tag{5}$$

Under steady economic conditions, we obtain[2]:

$$g_E = \frac{(\sigma - 1)}{(1 + \omega)} \frac{1}{\sigma} \left[\mu_2 \left(\frac{Y}{K} \right) - \rho \right] \tag{6}$$

where gE indicates variable change, and if it increases in the positive direction, the environment improves. Green credit ξ acts on the material capital of environmental protection enterprises, where a mean $K_l = \xi K$. Substituting production Equation (5) considering green credit into Equation (6), while taking the logarithms, we obtain Equations (7) and (8):

$$g_E = \frac{(\sigma - 1)}{(1 + \omega)} \frac{1}{\sigma} \left[\mu_2 h^{\alpha_3} A^{\alpha_2} N^{\mu_1} K_h^{\mu_2 - 1 - \alpha_2 \gamma} K_l^{\alpha_2 \gamma} (1 - \xi) - \rho \right] \tag{7}$$

$$\begin{aligned} \ln g_E = {} & \phi + \alpha_3 \ln h + (\alpha_1 \beta_1 - \alpha_3) \ln N + (\alpha_1 \beta_2 - 1) \ln K_h \\ & + \alpha_2 \gamma \ln K_l - \ln (1 - \xi) \end{aligned} \tag{8}$$

where g_E, h, N, K_h, K_l, and ξ are the proportion of environmental change, the environmental protection technology level, natural resource exploitation, the credit of polluting enterprises, the credit of cleaning enterprises, and the credit of polluting enterprises. A positive increase of g_E means that the environment is better. Equation (8) shows that the input of credit resources into green enterprises l can increase their output, while decreasing resource consumption and environmental pollution.

4. Method

4.1. Quantile regression model

This paper constructs a benchmark regression model of green credit on social health costs, as follows:

$$\begin{aligned} Health_{it} = {} & \alpha_0 + \alpha_1 Credit_{it} + \alpha_i Control_{it} + \lambda_i + \varepsilon_{it} i = 1, 2, \cdots, N \\ t = {} & 1, 2, \cdots, T \end{aligned} \tag{9}$$

where $Health_{it}$ is the social health cost, $Credit_{it}$ is green credit, $Control_{it}$ represents the control variable, α_1 is the effect of green credit on social health costs, λ_i represents the fixed effect, and ε_{it} represents the residual and follows a normal distribution.

When the benchmark regression Equation (9) is estimated, a traditional OLS can obtain only the impact of explanatory variables on the expected value of the explained variables. The OLS cannot analyze the influence of the distribution law. The QR method proposed by Koenker and Bassett (49) can solve this problem. The method assumes that the quantile of the conditional distribution of the dependent variable is a linear function of the independent variable, resulting in the construction of the QR of the dependent variable. Moreover, a QR can determine whether the independent variable has a specific regression curve on each characteristic quantile of the dependent variable. Therefore, compared with an OLS, a QR can more comprehensively describe the influence of the independent variables on the variation range of the dependent variables and show the conditional distribution shape. On the other hand, a QR uses the weighted average of the residual absolute value as the minimized objective function. Compared with an OLS, a QR is not easily affected by extreme values, and the estimation result is more robust.

Because the differences in economic development levels and policy implementation effects of regions in China, this paper uses a QR to analyze more comprehensively the impact of green credit on social health costs. The QR is defined as:

$$Quant_\theta \left(Health_{it} | X_{it} \right) = \alpha^\theta X_{it} \tag{10}$$

where X_{it} includes $Credit_{it}$ and $Control_{it}$, α^θ is the coefficient variable, and $Quant_\theta \left(Health_{it} | X_{it} \right)$ is the conditional quantile corresponding to quantile θ $(0 < \theta < 1)$of the social health cost given X. Coefficient vector α^θ corresponding to θ is realized by minimizing the absolute deviation. The estimated value of the regression coefficient should minimize the following objective functions:

$$\begin{aligned} \alpha^\theta = {} & \text{argmin} \left\{ \sum_{it, Health_{it} \geq X_{it}\alpha} \theta | Health_{it} - X_{it}\alpha \right. \\ & \left. + \sum_{it, Health_{it} \geq X_{it}\alpha} (1 - \theta) | Health_{it} - X_{it}\alpha \right\} \end{aligned} \tag{11}$$

Obviously, different regression lines are obtained for different θ. With a value of θ from 0 to 1, we obtain all the trajectories of the conditional distribution of explained variable $Health_{it}$, that is, a cluster of curves. Therefore, compared with the OLS mean regression with only one regression curve, the QR can more fully reflect the relationship between the model variables.

4.2. Mediation effect model

Assuming that the variables are continuous and standardized, the regression model considering environmental pollution (Env) as a mediating variable is as follows:

$$Env_{it} = \gamma_0 + \gamma_1 Credit_{it} + \gamma_i Control_{it} + \lambda_i + \nu_{it} \tag{12}$$

$$\begin{aligned} Health_{it} = {} & \beta_0 + \beta_1 Credit_{it} + \beta_2 Env_{it} + \beta_i Control_{it} + \lambda_i \\ & + \mu_{it} \end{aligned} \tag{13}$$

In Equation (12), γ_1 is the effect of green credit on environmental pollution. In Equation (13), β_2 is the effect of mediating variable

2 See Appendix A for details.

environmental pollution on social health costs after controlling for green credit. λ_i represents the fixed effect, and μ_{it} *and* ν_{it} represent the residuals, assuming that they all follow a normal distribution and are independent of one another. Substituting Equation (13) into Equation (12), we obtain:

$$Health_{it} = k_0 + (\beta_1 + \gamma_1 \cdot \beta_2)\, Credit_{it} + k_i Control_{it} + \nu_{it} \cdot \beta_2 + \mu_{it} \quad (14)$$

where $\gamma_1 \cdot \beta_2$ is the mediating effect of green credit on social health costs, β_1 is the direct effect of green credit on social health costs, and $\beta_1 + \gamma_1 \cdot \beta_2$ is the total effect of green credit on social health costs. It can be concluded that $\alpha_1 = \beta_1 + \gamma_1 \cdot \beta_2$. Therefore, the mediating effect of environmental pollution can be estimated to observe its role in the social health costs of green credit.

5. Data

This paper is based on the data of 30 administrative regions of mainland Chinese provinces for 2005–2020. Tibet was excluded because of a lack of data. The data sources are the China Statistical Yearbook, China Environmental Statistical Yearbook, China Urban Statistical Yearbook, and Economic Census Yearbook. In 2005, BankTrack investigated the implementation of the Equator Principles[3] and thought that some institutional members were just "green washing." This phrase means that some enterprises do not consider environmental factors in their daily business activities and investment decisions but still aim to maximize their profits and is a phenomenon that has aroused worldwide attention. In 2007, China issued the Opinions on Implementing Environmental Protection Policies and Regulations to Prevent Credit Risk, which noted that green credit, as an economic means, has become a primary form of pollution reduction. In 2012, the China Banking Regulatory Commission (CBRC) issued Guidelines on Green Credit, which put forward clear requirements for banking financial institutions to engage in green credit and to vigorously promote energy conservation, emissions reductions, and environmental protection. In 2022, according to the CBRC, the balance of green credit at the end of the third quarter increased by 29.5% compared with the beginning of the year. With the orderly promotion of the banking industry, green credit developed rapidly. At present, there are three main ways to measure the green credit scale: the dummy variables of green credit policies (28), bank loans aimed at energy conservation and environmental protection (3), and the credit share of energy-intensive and highly polluting industries. As the dummy variables are more suitable for comparing green credit policies, and this paper uses provincial level data, the credit share can indirectly and partially reflect the regional green credit development level (17, 50). The credit share refers to the ratio of the interest expenditures of the six high energy–consuming industries[4] to the industrial output. We used the negative value of this ratio to measure green credit. Green credit can affect financial funds or residents' health in a variety of ways, while the deterioration of the environment threatens residents' health, so more medical financial expenditures are needed. We chose per capita financial expenditures on medical care to measure social health costs. Through a theoretical analysis, environmental pollution, as a mediating variable, has been found to play a substantial role in the model of green credit affecting health. Based on the industrial wastewater, industrial SO_2, and industrial dust discharges, we used the entropy method to calculate the comprehensive environmental index.

Among the control variables, we selected three presentation modes to govern the environment: the treatment number of industrial waste gas (Gas) and water facilities (Water) and the comprehensive utilization amount of industrial solid waste (Solid). The financial expenditure on environmental governance (Fee) represents various levels of regional support for environmental protection. The health status is affected by population aging. Generally, places with a large number of older adult people (Old) require more medical and health expenditures, so we chose the older adult dependency ratio to measure the aging degree. In addition, the economic development level is an important factor affecting health expenditures and environmental pollution. We chose GDP per capita and took 2005 as the base period for the decrease.

The variables' descriptive statistics are shown in Table 1. The average value of social health costs were 24,900 yuan/person, with a large difference between the extreme values (8.622). The average level of green credit in various regions was 0.512, and the environmental pollution level was 0.539, with little difference between the extreme values. However, except for the variables old and GDP, the extreme values of the control variables differed greatly, a fact that can also be seen from their standard deviations.

Figure 1A shows the trend in national green credit and social health costs from 2005 to 2020. The two variables developed in different directions. Social health costs increased over time, from 0.01 in 2005 to 0.148 in 2020[5], an increase of 14%, with an average annual increase of 0.875%. This result shows that China's financial investment in social health expenditures has increased, especially since COVID-19 in 2019, and the trend is obviously increasing, which introduces great challenges to finance. Green credit fluctuated greatly during the review period. It increased from 2005–2009 and declined significantly from 2009–2013, from 0.516 to 0.474, for a decline of 8.94%. It increased slightly in 2013–2014 and then began to decrease sharply, rebounding until 2016 and reaching 0.479 in 2020. However, this figure was still far below the peak of 0.516, reached in 2009.

To observe the green credit and social health costs in different regions, according to the economic geographical location, the whole sample is divided into the Eastern, Central, and Western regions[6].

3 According to the policy guidelines of the International Finance Corporation and the World Bank, the financial industry has established benchmarks aimed at judging, assessing, and managing environmental or social risks in project financing to increase the social responsibility of the banking industry and improve the increasingly stressed environment.

4 The six high-energy consuming industries include petroleum processing, coking and nuclear fuel processing, chemical raw materials and chemical products manufacturing, non-metallic mineral products, ferrous metal smelting and rolling processing, nonferrous metal smelting and rolling processing, and the production and supply of electricity and heat.

5 The health data is based on the secondary axis.

6 Eastern region: Liaoning, Beijing, Tianjin, Shanghai, Hebei, Shandong, Jiangsu, Zhejiang, Fujian, Guangdong, Guangxi, and Hainan; Central region: Heilongjiang, Jilin, Shanxi, Inner Mongolia, Anhui, Henan, Huzibei, Hunan, and Jiangxi; and Western region: Chongqing, Sichuan, Yunnan, Guizhou, Tibet, Shaanxi, Gansu, Qinghai, and Ningxia.

TABLE 1 Variable descriptive statistics.

	Variable	Unit	Numbers	Mean	Std.	Min.	Max.
Dependent variable	Health	10,000 yuan/person	480	0.249	0.849	0.002	8.622
Independent variable	Credit	ratio	480	0.512	0.203	0.010	2.411
Mediating variable	Env	–	480	0.539	0.530	0.001	2.585
Control variable	Gas	set	480	8508.719	7932.646	332.000	57278.000
	Water	set	480	2586.198	2250.746	103.000	10608.000
	Solid	10,000 tons	480	5794.541	4778.710	87.000	25230.000
	Fee	10,000 yuan/person	480	375.481	1716.833	0.551	37432.540
	Old	ratio	480	13.778	3.360	7.440	25.480
	GDP	10,000 yuan/person	480	4.421	2.785	0.531	16.493

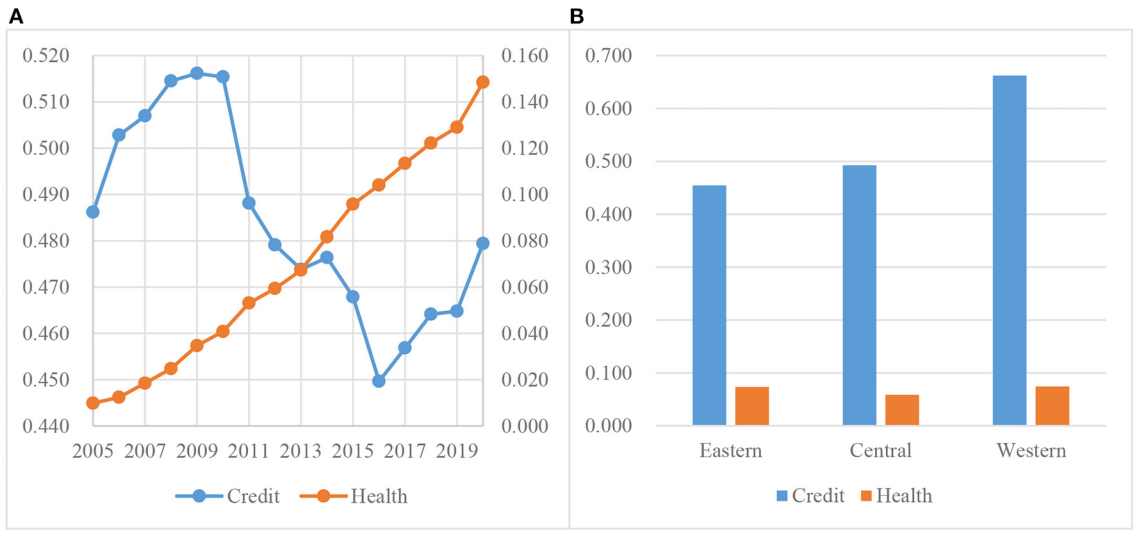


FIGURE 1
(A, B) Trend and regional difference in credit and health.

Observing the regional differences in green credit in Figure 1B, we found that the financial institutions in the Eastern region released the least amount of green credit, followed by the Central region and the Western region. This result may be due to the continual migration of polluting enterprises from the Eastern region to inland areas. Most of such enterprises are high-tech industries such as tertiary ones. Regarding the regional differences in social health costs, we found that the Central region had the least financial expenditures on health care, while the difference between the Eastern and Western regions was practically negligible. This result shows that the fiscal expenditures on health in the Central region lag behind those in the Eastern and Western regions, but the difference is not significant.

6. Empirical analysis

6.1. Quantile regression estimation results

To determine the function form, the Hausman test result for Equation (9) was 26.55, and the *p*-value was 0.00. Therefore, it is appropriate to reject the original assumption that individuals are random effects and to select a fixed effect model. For comparison with the panel QR results, we also show a fixed effect model estimation (Column 7). In the panel quantile estimation (Table 2), the five representative quantiles of 10, 25, 50, 75, and 90% are selected.

The elasticity coefficient of green credit was negative in both the OLS regression and each quantile level, all of which except the 10% quantile level passed the 10% significance level test. Since 2005, the green credit of various institutions in China decreased social health costs and relieved financing pressures on health expenditures. The change trend of each quantile reveals that as the quantile increased, the absolute value of the green credit coefficient gradually increased from 0.162 to 0.641. This change shows that for provinces with high per capita financial health expenditures, green credit plays a greater role in decreasing social costs (51, 52). Green credit plays an indispensable role in decreasing financial expenditures and protecting public health. For example, the financial industry quickly took measures to adjust credit to prevent and control COVID-19 and supported various industries as they attempted to resume production. The main methods include deferred repayment, the reduction of late

TABLE 2 Quantile regression results of green credit on social health costs.

	10%	25%	50%	75%	90%	FE
Credit	−0.162 [−0.9]	−0.240* [−1.78]	−0.338*** [−2.97]	−0.443*** [−2.93]	−0.641** [−2.19]	−0.347*** [−3.23]
Gas	−0.811*** [−3.32]	−0.615*** [−3.43]	−0.365*** [−2.37]	−0.100 [−0.49]	0.399 [1.02]	−0.344*** [−2.66]
Water	−0.004 [−0.02]	−0.146 [−1.08]	−0.327*** [−2.80]	−0.519*** [−3.39]	−0.881*** [−2.96]	−0.343*** [−3.39]
Solid	0.147 [1.15]	0.034 [0.37]	−0.108 [−1.34]	−0.260*** [−2.47]	−0.546*** [−2.66]	−0.121* [−1.93]
Fee	−0.170 [−1.16]	−0.104 [−0.95]	−0.019 [−0.21]	0.071 [0.57]	0.239 [1.01]	−0.012 [−0.16]
GDP	0.770*** [3.99]	0.789*** [5.45]	0.812*** [6.66]	0.837*** [5.15]	0.884*** [2.82]	0.814*** [8.39]
Old	−0.674** [−1.89]	−0.539** [−2.02]	−0.367 [−1.63]	−0.184 [−0.62]	0.159 [0.28]	−0.352* [−1.72]
Intercept	–	–	–	–	–	3.471*** [4.35]

*, **, ***represent 10, 5, 1% significance levels, respectively.

fees, the reduction of interest rates for small and micro enterprises, and the establishment of special funds to match targets. In early 2020, China Development Bank launched the CDB Anti -epidemic Special Bond to effectively provide financial support to fight the epidemic. All the above measures can decrease social health expenditures. The literature also supports the conclusion of this paper by affirming the important contribution of green credit in decreasing public health expenditures (10, 11). In addition, Hu et al. (53) used a QR to show that the increasing effect of green credit on green total factor productivity increased as the latter increased.

The QR results of each control variable show that the first 50% quantile, the elasticity coefficient of industrial waste gas was significantly negative. Industrial wastewater was negative in all fractions, but only the last 50% of the quantiles passed the 10% significance level test, and industrial solid waste was significantly negative in the last 75% quantile. Although the significance of industrial waste differs, in general, the treatment of industrial wastewater, waste gas and solid waste can have a positive effect on decreasing social health costs. The elasticity coefficient of fiscal environmental expenditures in the first 25% quantile was negative but not significant. The elasticity coefficient of GDP was positive in all quantiles, passing the 10% significance level. This result indicates that regions with high economic development levels and high also have relatively high per capita medical and health expenditures. As the quantile level increased, the elasticity coefficient increased from 0.770 to 0.884. This result means that economic growth has a greater role in decreasing social health costs in regions where such costs are high. The elasticity coefficient of population aging was significantly negative in the first 50% quantile. This result indicates that as population aging intensifies, per capita medical health expenditures did not increase. The impact is limited of aging on the increase in medical and health expenditures, especially in rural areas. Because economic development is at a low level, the demand for medical services for the older adult population has not been effectively met due to the imperfect medical security system and the lack of medical resources. Therefore, at this stage, rural aging has not affected medical expenditures, and even as the aging degree improves, self-funded medical health expenditures have decreased.

In the fixed effect regression model, the regression coefficient of green credit was −0.347, which was significantly negative at the 1% level, indicating that implementing green credit can decrease social health costs. The elasticity coefficient of each control variable was also numerically equivalent to the mean QR value, which verifies the robustness of the results.

6.2. Regional heterogeneity analysis

This paper divides the whole sample into three regions—Eastern, Central, and Western—and comparatively analyzes the regional differences. Table 3 lists only the elasticity coefficients of green credit under different quantiles and fixed effects, absent the control variables.

Table 3 shows that the fixed effect coefficient was basically maintained at the average level of each quantile and thus verifies the robustness of the QR results. The elasticity coefficient of green credit in the Eastern region was negative in all quantiles. As the quantile increased, the absolute value of the elasticity coefficient increased from 0.557 to 0.854, all of which passed the 1% significance level test. This result verifies that the increase of green credit can decrease social health costs, and the effect is greater and more significant than is that under the full sample condition. The elasticity coefficients of green credit in the Central region were all significantly negative (except for the low quantile of 10%). Furthermore, as the quantile increased, the absolute value increased from 0.270 to 0.379. This result verifies the negative relationship between green credit and social health costs, but the negative effect is lower than was that in the Eastern region. Except for the 90% high quantile in the Western region, the elasticity coefficients of the quantiles were positive and not significant, indicating that no effect was found for green credit decreasing social health costs in the Western region. In short, the impact of green credit on social health costs shows

TABLE 3 Regional heterogeneity.

	Eastern	Central	Western
10%	−0.557*** [−2.77]	−0.270 [−1.19]	0.337 [0.71]
25%	−0.621*** [−4.24]	−0.295* [−1.79]	0.209 [0.73]
50%	−0.698*** [−5.85]	−0.322** [−2.33]	0.107 [0.58]
75%	−0.797*** [−4.62]	−0.363* [−1.70]	0.024 [0.11]
90%	−0.854*** [−3.76]	−0.379* [−1.83]	−0.032 [−0.12]
FE	−0.702*** [−5.42]	−0.326** [−2.07]	0.125 [0.62]
Samples number	176	128	176

*, **, ***represent 10, 5, 1% significance levels, respectively.

TABLE 4 Bootstrap test results.

	Coef.	BootSE	LLCI	ULCI
Direct effect	0.011	0.012	−0.012	0.053
Indirect effect	−0.864***	0.092	−1.045	−0.683

*, **, ***represent 10, 5, 1% significance levels, respectively; LLCI and ULCI are the lowest and highest values of the confidence interval; BootSE is the standard error.

TABLE 5 Regression results of the mediating effect.

	Dependent variable: Env	Dependent variable: Health
Credit	−0.969*** [−3.40]	−0.298** [2.14]
Env	–	0.183*** [2.86]
Intercept	1.405 [1.37]	5.747** [1.99]
Control	Yes	Yes
F	18.41	17.37
Adj-*R*2	0.572	0.684

*, **, ***represent 10, 5, 1% significance levels, respectively.

an obvious geographical ladder distribution, with the largest effect in the Eastern region, followed by the Central region, and an as-yet-to-be-found effect in the Western region. This conclusion is consistent with Zhong's (31) view that the development of digital finance in the Eastern region is generally ahead of other regions, and the environmental improvement is more effective. Li et al. (7) also believed that green credit improved the total factor productivity only in the Eastern region but had little impact in other regions, so the promotional effect of financial and legal developed areas would be more effective. Therefore, the Chinese government should encourage the implementation of regionally differentiated green credit policies, increase investment in less-developed regions, and pay attention to the efficiency of capital use.

6.3. Verification of the mediating effect of the environment

In the process of decreasing social health costs, green credit inevitably considers the important role played by environmental pollution. Most researchers recommend using the bootstrap method for a mediating effect analysis (54, 55) as it corrects the interval estimation error by adjusting the percentile of the sequence interval. Regarding H0, $\gamma_1 \cdot \beta_2 = 0$, when the sample size is 5,000, the results of the non-parametric percentile bootstrap method are shown in Table 4. Under the 95% confidence interval, the intermediate test results did not include 0 (LLCI = −1.045, ULCI = −0.683), so H0 is rejected, indicating that the mediating effect of environmental pollution was significant, and the size of the mediating effect was −0.864. In addition, after controlling for the mediating variable environmental pollution, green credit did not have a significant impact on social health costs, and the confidence interval included 0 (LLCI −0.012, ULCI = 0.053). Therefore, environmental pollution plays a mediating role in the pathway of green credit to social health cost and is the only mediating variable.

The estimated results for Equations (12) and (13) are shown in Table 5, which examines the mediating effect of environmental pollution. Equation (12) takes environmental pollution as the explained variable. The elasticity coefficient of green credit was −0.269, passing the 1% significance level test, indicating that green credit plays an important role in environmental protection. On the one hand, through green credit tools, the banking industry has increased support for carbon emissions trading and environmental liability insurance, which have alleviated human damage to the environment. On the other hand, green credit can effectively encourage enterprises to carry out green technology innovation. For example, enterprises widely use clean energy, clean materials, and green technologies and processes to replace their original ones, ultimately decreasing the risk of environmental pollution. Therefore, green credit is a financial activity that supports environmental improvement, responds to climate change, and effectively uses resources and is an important guarantee for sustainable economic development.

The setting of Equation (13) in Table 5 takes social health cost as the explained variable. The elasticity coefficient of the green credit was −0.298, while environmental pollution was 0.183, which were all significant at the 1% level. This result not only supports the role of green credit in decreasing health costs but also shows that the increase of industrial waste emissions might increase the social expenditures on health. We analyze the conclusion from both psychological and physical aspects. On the one hand, environmental pollution has a negative effect on public mental health, decreasing subjective wellbeing and mental health (56). On the other hand, environmental pollution accelerates the depreciation of physical health, decreases the productivity of personal exercise used

TABLE 6 Quantile regression results of green credit on health care expenditures.

	10%	25%	50%	75%	90%	FE
Credit	−0.236*** [−3.86]	−0.574*** [−6.92]	−1.019*** [−11.5]	−1.463*** [−10.71]	−1.854*** [−8.52]	−1.006*** [−15.61]
Gas	−0.565*** [−2.66]	−0.464*** [−3.06]	−0.332*** [−2.92]	−0.200 [−1.30]	−0.083 [−0.37]	−0.352*** [−3.33]
Water	−0.508*** [−2.73]	−0.573*** [−4.31]	−0.658*** [−6.61]	−0.743*** [−5.51]	−0.817*** [−4.14]	−0.644*** [−7.63]
Solid	0.291*** [2.70]	0.275*** [3.56]	0.254*** [4.42]	0.233*** [2.98]	0.215* [1.87]	0.270*** [4.75]
Fee	−0.075 [−0.59]	−0.072 [−0.78]	−0.067 [−0.98]	−0.062 [−0.67]	−0.058 [−0.42]	−0.029 [−0.49]
GDP	−0.050 [−0.22]	−0.149 [−0.91]	−0.278** [−2.27]	−0.407** [−2.46]	−0.521** [−2.15]	−0.248** [−2.32]
Old	−0.587 [−1.57]	−0.682** [−2.54]	−0.807*** [−4.03]	−0.931*** [−3.42]	−1.040*** [−2.61]	−0.773*** [−4.59]
Intercept	–	–	–	–	–	10.338*** [11.97]

*, **, ***represent 10%, 5%, 1% significance levels, respectively.

to increase health investments, and negatively affects the public's physical health. Guo et al. (57) proved that industrial structure upgrades and environmental investment play positive intermediary roles between green credit and green economy efficiency. Moreover, Zeng et al. (17) and Muhammad et al. (18) supported the research results using different national data sources and methods.

In summary, green credit has a direct effect on social health costs if −0.298, an indirect effect of −0.049 (=-0.269*0.183), and a total effect for −0.347 [=-0.298+(−0.269*0.183)]. The calculated total effect is consistent with the elasticity coefficient of green credit in the overall regression, Equation (9) in Table 2 (Column 7). This result shows that green credit directly impacts social health costs and indirectly affects residents' health through environmental pollution.

6.4. Robustness check

Table 2 lists the fixed effect regression results. The sign of the green credit in the fixed effect model was consistent with the QR, and the size was basically the same as it was in the average quantile, which indicates the robustness of the QR model. In this section, we measure the explained variable social health cost by per capita medical care expenditures to further test the robustness of this research object.

In Table 6, after the explained variable was replaced, the coefficient of green credit did not change much compared with the result in Table 2. Only the size or significance slightly decreased or improved, a fact that does not affect the conclusion of this paper. The elasticity coefficient of green credit was negative at the 10% significant level, which further indicates that green credit can decrease social health expenditures.

7. Conclusion

Green credit is an innovative financial concept and reflects the sustainable development of an economy and a society. Based on data from 30 provinces in China from 2005 to 2020, this paper examines the different impacts of green credit on social health costs using a QR. In addition, we examine the mediating role of environmental pollution in the impact path and mainly draw the following conclusions. The implementation of green credit policies can decrease social health costs. As the quantile improves, the absolute value of the green credit coefficient gradually increases. This change shows that as per capita financial health expenditures increase, green credit plays a greater role in decreasing social costs. This effect is strongest in the Eastern region and may be related to the economic development level. In addition, this paper confirms the mediating effect of environmental pollution and the fact that it is the only mediating variable, which has also been proven by the green credit and health general equilibrium model. Therefore, green credit can relieve financial pressure through social financing, release more funds for public health expenditures, and affect public health by improving the environmental quality.

The government should make full use of financial supports to help build a human health community and should improve the emergency management capabilities for major public health events as soon as possible. Policy-makers need to fully support internet green finance led by green credit, encourage the diversified development of green credit products, and increase the promotion of green credit products. At the same time, banking institutions should strengthen the positive innovation of green credit products, expand the modes of public participation, respond to the rapidly changing market, and decrease government pressure. In addition, green finance is the embodiment of the green development concept and can protect the human living environment and decrease the impact of pollution on residents' health (58). Therefore, financial institutions can develop personalized green credit products for specific pollutant types or industries. These practices will help to improve financial institutions' efficiency as they undertake environmental responsibilities and decrease costs.

8. Limitations

The impact of green finance on residents' health is not limited to green credit. Because green credit accounts for a

large proportion of green finance, the scope of this paper is green credit. As green stocks, bonds, trusts, etc. gradually develop green finance's research objects should be gradually expanded to more comprehensively analyze green finance's impact on residents' health.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YR: conceptualization, methodology, software, data curation, and writing—original draft preparation. JH: visualization and investigation and writing—reviewing and editing. Both authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1121154/full#supplementary-material

References

1. Cheng S, Xiang Z, Xi H. Environmental status and human health: evidence from China. *Int J Environ Res Public Health.* (2022) 19:12623. doi: 10.3390/ijerph191912623.

2. Liu Q, Dong B. How does China's green credit policy affect the green innovation of heavily polluting enterprises? The perspective of substantive and strategic innovations. *Environ Sci Pollut Res.* (2022) 29:77113–30. doi: 10.1007/s11356-022-21199-6.

3. Wang Y, Lei X, Zhao D, Long R, Wu M. The dual impacts of green credit on economy and environment: evidence from China. *Sustainability.* (2021) 13:4574. doi: 10.3390/su13084574

4. Zhao J, Wang J, Dong K. The role of green finance in eradicating energy poverty: Ways to realize green economic recovery in the post-COVID-19 era. *Econ Change Restruct.* (2022). doi: 10.1007/s10644-022-09411-6

5. Murray CJ, Lopez AD. Measuring the global burden of disease. *New England J Med.* (2013) 369:448–57. doi: 10.1056/NEJMra1201534

6. Landrigan PJ, Fuller R, Acosta NJR, Adeyi O, Arnold R, Basu N, et al. The lancet commission on pollution and health. *Lancet.* (2018) 391:462–512. doi: 10.1016/s0140-6736(17)32345-0

7. Li B, Zhang J, Shen Y, Du Q. Can green credit policy promote green total factor productivity? Evidence from China environmental science and pollution. *Research.* (2022) 3:5. doi: 10.1007/s11356-022-22695-5

8. Federspiel F, Borghi J, Martinez-Alvarez M. Growing debt burden in low- and middle-income countries during COVID-19 may constrain health financing. *Glob Health Action.* (2022) 15:2461. doi: 10.1080/16549716.2022.2072461

9. Tuohy CH, Flood CM, Stabile M. How does private finance affect public health care systems? Marshaling the evidence from OECD nations. *J Health Polit Policy Law.* (2004) 29:359–96. doi: 10.1215/03616878-29-3-359

10. Leatherman S, Geissler K, Gray B, Gash M. Health financing: a new role for microfinance institutions? *J Int Dev.* (2013) 25:881–96. doi: 10.1002/jid.2829

11. Maurya D, Asher M. Sustainable health financing system for India: the economic perspective. *Nat Med J India.* (2021) 34:95–9. doi: 10.4103/0970-258x.326759.

12. Liu X, Guo SQ. Inclusive finance, environmental regulation, and public health in China: lessons for the COVID-19 pandemic. *Front Public Health.* (2021) 9:2166. doi: 10.3389/fpubh.2021.662166

13. Hu YQ, Jiang HY, Zhong ZQ. Impact of green credit on industrial structure in China: theoretical mechanism and empirical analysis. *Environ Sci Pollut Res.* (2020) 27:10506–19. doi: 10.1007/s11356-020-07717-4

14. Zhu XW. Does green credit promote industrial upgrading? Analysis of mediating effects based on technological innovation. *Environ Sci Pollut Res.* (2022) 29:41577–89. doi: 10.1007/s11356-021-17248-1

15. Qin JH, Cao JH. Carbon emission reduction effects of green credit policies: empirical evidence from China. *Front Environ Sci.* (2022) 10:8072. doi: 10.3389/fenvs.2022.798072

16. Wang XX, Huang JY, Xiang ZM, Huang JL. Nexus between green finance, energy efficiency, and carbon emission: COVID-19 implications from BRICS countries. *Front Energy Res.* (2021) 9:6659. doi: 10.3389/fenrg.2021.786659

17. Zeng HL, Iqbal W, Chau KY, Shah SAR, Ahmad W, Hua H. Green finance, renewable energy investment, and environmental protection: empirical evidence from BRICS countries. *Econ Res-Ekonomska Istrazivanja.* (2021). doi: 10.1080/1331677x.2022.2125032

18. Sadiq M, Amayri MA, Paramaiah C, Mai NH, Ngo TQ, Phan TTH. How green finance and financial development promote green economic growth: deployment of clean energy sources in South Asia. *Environ Sci Pollut Res.* (2022) 29:65521–34. doi: 10.1007/s11356-022-19947-9

19. Zhang KQ, Chen HH, Tang LZ, Qiao S. Green finance, innovation and the energy-environment-climate nexus. *Front Environ Sci.* (2022) 10:9681. doi: 10.3389/fenvs.2022.879681

20. Qiao HS, Zhu J, Huang H, editors. The impact of financial development on environment: An empirical analysis from 1996 to 2005 in China. *International Conference on Information Technology and Industrial Engineering 2013 Aug 07–08.* Wuhan, Peoples R China (2014).

21. Sun JX, Wang F, Yin HT, Zhang B. Money talks: the environmental impact of China's green credit policy. *J Policy Anal Manag.* (2019) 38:653. doi: 10.1002/pam.22137

22. Huang YM, Chen C, Lei LJ, Zhang YP. Impacts of green finance on green innovation: a spatial and nonlinear perspective. *J Clean Prod.* (2022) 365:2548. doi: 10.1016/j.jclepro.2022.132548

23. Tang DC, Zhong H, Zhang JY, Dai YG, Boamah V. The effect of green finance on the ecological and environmental quality of the Yangtze River economic belt. *Int J Environ Res Public Health.* (2022) 19:2492. doi: 10.3390/ijerph191912492

24. Xu SY, Zhu HQ. Does green governance efficiency and green finance policies matters in sustainable environment: implications for public health. *Front Public Health.* (2022) 10:1349. doi: 10.3389/fpubh.2022.861349

25. Liu S, Xu RX, Chen XY. Does green credit affect the green innovation performance of high-polluting and energy-intensive enterprises? Evidence from a quasi-natural experiment. *Environ Sci Pollut Res.* (2021) 28:65265–77. doi: 10.1007/s11356-021-15217-2

26. Chen ZG, Zhang YQ, Wang HS, Ouyang X, Xie YX. Can green credit policy promote low-carbon technology innovation? *J Clean Prod.* (2022) 359:2061. doi: 10.1016/j.jclepro.2022.132061

27. Hong M, Li ZH, Drakeford B. Do the green credit guidelines affect corporate green technology innovation? Empirical research from China. *Int J Environ Res Public Health*. (2021) 18:1680. doi: 10.3390/ijerph18041682

28. Yang Y, Zhang YL. The Impact of the green credit policy on the short-term and long-term debt financing of heavily polluting enterprises: based on PSM-DID Method. *Int J Environ Res Public Health*. (2022) 19:1287. doi: 10.3390/ijerph191811287

29. Chai SL, Zhang K, Wei W, Ma WY, Abedin MZ. The impact of green credit policy on enterprises' financing behavior: evidence from Chinese heavily-polluting listed companies. *J Clean Prod*. (2022) 363:2458. doi: 10.1016/j.jclepro.2022.132458

30. Zhou XG, Tang XM, Zhang R. Impact of green finance on economic development and environmental quality: a study based on provincial panel data from China. *Environ Sci Pollut Res*. (2020) 27:19915–32. doi: 10.1007/s11356-020-08383-2.

31. Zhong KY. Does the digital finance revolution validate the environmental Kuznets curve? Empirical findings from China. *PLoS ONE*. (2022) 17:7498. doi: 10.1371/journal.pone.0257498

32. Grossman M. On the concept of health capital and the demand for health. *J Polit Econ*. (1972) 80:223–55. doi: 10.1086/259880

33. Cropper ML. Measuring the benefits from reduced morbidity. *American Economic Review*. (1981) 71:235–40.

34. Wagstaff A. The demand for health—Some new empirical evidence. *J Health Econ*. (1986) 5:195–233. doi: 10.1016/0167-6296(86)90015-9

35. Scally G, Perkins C. Environment and health. *Hospital Med*. (1998) 59:872–76.

36. Welsch H. Preferences over prosperity and pollution: environmental valuation based on happiness surveys. *Kyklos*. (2002) 55:473–94. doi: 10.1111/1467-6435.00198

37. Knibbs LD, de Waterman AMC, Toelle BG, Guo YM, Denison L, Jalaludin B, et al. The Australian child health and air pollution study. (ACHAPS): a national population-based cross-sectional study of long-term exposure to outdoor air pollution, asthma, and lung function. *Environ Int*. (2018) 120:394–403. doi: 10.1016/j.envint.2018.08.025

38. Piracha A, Chaudhary MT. Urban air pollution, urban heat island and human health: a review of the literature. *Sustainability*. (2022) 14:9234. doi: 10.3390/su14159234

39. Levinson A. Valuing public goods using happiness data: the case of air quality. *J Public Econ*. (2012) 96:869–80. doi: 10.1016/j.jpubeco.2012.06.007

40. Liu L, Wang K, Wang SS, Zhang RQ, Tang XY. Assessing energy consumption, co2, and pollutant emissions and health benefits from China's transport sector through 2050. *Energy Policy*. (2018) 116:382–96. doi: 10.1016/j.enpol.2018.02.019

41. Ljungman PLS Li WY, Rice MB, Wilker EH, Schwartz J, Gold DR, et al. Long- and short-term air pollution exposure and measures of arterial stiffness in the Framingham heart study. *Environ Int*. (2018) 121:139–47. doi: 10.1016/j.envint.2018.08.060

42. Alberini A, Cropper M, Fu TT, Krupnick A, Liu JT, Shaw D, et al. Valuing health effects of air pollution in developing countries: the case of Taiwan. *J Environ Econ Manage*. (1997) 34:107–26. doi: 10.1006/jeem.1997.1007

43. Ivanova VR. The anthropogenic air pollution and human health. *J Imab*. (2020) 26:3057–62. doi: 10.5272/jimab.2020262.3057

44. Zhao JK, Yang WP, Zho N, Ain QU, Zhao K. Measurement and space-time evolution of health level under constraint of environmental pollution, China: 2002–2016. *Environ Sci Pollut Res*. (2020) 27:26725–41. doi: 10.1007/s11356-020-08931-w

45. Wang SJ, Zhou HY, Hua GH, Wu Q. What is the relationship among environmental pollution, environmental behavior, and public health in China? A study based on CGSS. *Environ Sci Pollut Res*. (2021) 28:20299–312. doi: 10.1007/s11356-020-11951-1

46. Wang Q, Li LZ, Zhang YP, Cui Q, Fu YZ, Shi WY, et al. Research on the establishment and application of the environmental health indicator system of atmospheric pollution in China. *Bull Environ Contam Toxicol*. (2021) 106:225–34. doi: 10.1007/s00128-020-03084-5

47. Chen Z, Wang JN, Ma GX, Zhang YS. China tackles the health effects of air pollution. *Lancet*. (2013) 382:1959–60. doi: 10.1016/s0140-6736(13)62064-4

48. Bovenberg AL, Smulders S. Environmental-quality and pollution-augmenting technological-change in a 2-sector endogenous growth-model. *J Public Econ*. (1995) 57:369–91

49. Koenker R, Bassett G Jr. Regression quantiles. *Econ J Econ Soc*. (1978) 3:33–50. doi: 10.2307/1913643

50. Lin BQPT. How does environmental regulation affect the development of green credit? *China Populat Res Environ*. (2022) 32:11.

51. ul Husnain MI, Beyene SD, Aruga K. Investigating the energy–environmental Kuznets curve under panel quantile regression: a global perspective. *Environ Sci Pollut Res*. (2022) 3:3. doi: 10.1007/s11356-022-23542-3

52. Raghutla C, Padmagirisan P, Sakthivel P, Chittedi KR, Mishra S. The effect of renewable energy consumption on ecological footprint in N-11 countries: evidence from panel quantile regression approach. *Renewable Energy*. (2022) 197:125–37. doi: 10.1016/j.renene.2022.07.100

53. Hu QQ, Li X, Feng YH. Do green credit affect green total factor productivity? Empirical evidence from China. *Front Energy Res*. (2022) 9:1242. doi: 10.3389/fenrg.2021.821242

54. Preacher KJ, Hayes AF, SPSS. and SAS procedures for estimating indirect effects in simple mediation models. *Behav Res Meth Instrum Comput*. (2004) 36:717–31. doi: 10.3758/bf03206553

55. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr*. (2009) 76:408–20. doi: 10.1080/03637750903310360

56. Li ZT, Folmer H, Xue JH. To what extent does air pollution affect happiness? The case of the Jinchuan mining area. *China Ecol Econ*. (2014) 99:88–99. doi: 10.1016/j.ecolecon.2013.12.014

57. Guo LJ, Tan WY, Xu Y. Impact of green credit on green economy efficiency in China. *Environ Sci Pollut Res*. (2022) 29:35124–37. doi: 10.1007/s11356-021-18444-9

58. Lian YH, Gao JY, Ye T. How does green credit affect the financial performance of commercial banks? Evidence from China. *J Clean Prod*. (2022) 344:1069. doi: 10.1016/j.jclepro.2022.131069

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 02 February 2023
DOI 10.3389/fpubh.2023.1120462

OPEN ACCESS

EDITED BY
Ke Yan,
National University of Singapore, Singapore

REVIEWED BY
Kewen Su,
Hangzhou Hospital for the Prevention and Treatment of Occupational Disease, China
Qingfeng He,
Fudan University, China
Jian Wang,
University of Electronic Science and Technology of China, China

*CORRESPONDENCE
Yuan Yuan
✉ cq_double_yy@163.com
Xiaoxin Wu
✉ xiaoxinwu@zju.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

RECEIVED 10 December 2022
ACCEPTED 18 January 2023
PUBLISHED 02 February 2023

CITATION
Zhang Z, Tan J, Jin W, Qian H, Wang L, Zhou H, Yuan Y and Wu X (2023) Severe fever with thrombocytopenia syndrome virus trends and hotspots in clinical research: A bibliometric analysis of global research.
Front. Public Health 11:1120462.
doi: 10.3389/fpubh.2023.1120462

Severe fever with thrombocytopenia syndrome virus trends and hotspots in clinical research: A bibliometric analysis of global research

Zhengyu Zhang[1†], Juntao Tan[2†], Wen Jin[1], Hong Qian[3], Loulei Wang[1], Hu Zhou[4], Yuan Yuan[5*] and Xiaoxin Wu[6*]

[1]Medical Records Department, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2]Operation Management Office, Affiliated Banan Hospital of Chongqing Medical University, Chongqing, China, [3]Medical Records Department, The First Hospital of Lanzhou University, Lanzhou, China, [4]General Committee Office, The People's Hospital of Yubei District of Chongqing City, Chongqing, China, [5]Medical Department, Women and Children's Hospital of Chongqing Medical University, Chongqing, China, [6]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Centre for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Background: Since severe fever with thrombocytopenia syndrome virus (SFTSV) was first reported in 2009, a large number of relevant studies have been published. However, no bibliometrics analysis has been conducted on the literature focusing on SFTSV. This study aims to evaluate the research hotspots and future development trends of SFTSV research through bibliometric analysis, and to provide a new perspective and reference for future SFTSV research and the prevention of SFTSV.

Methods: We retrieved global publications on SFTSV from the Web of Science Core Collection (WoSCC) and Scopus databases from inception of the database until 2022 using VOSviewer software and CiteSpace was used for bibliometric analysis.

Results: The number of SFTSV-related publications has increased rapidly since 2011, peaking in 2021. A total of 45 countries/regions have published relevant publications, with China topping the list with 359. The Viruses-Basel has published the most papers on SFTSV. In addition, Yu et al. have made the greatest contribution to SFTSV research, with their published paper being the most frequently cited. The most popular SFTSV study topics included: (1) pathogenesis and symptoms, (2) characteristics of the virus and infected patients, and (3) transmission mechanism and risk factors for SFTSV.

Conclusions: In this study, we provide a detailed description of the research developments in SFTSV since its discovery and summarize the SFTSV research trends. SFTSV research is in a phase of explosive development, and a large number of publications have been published in the past decade. There is a lack of collaboration between countries and institutions, and international collaboration and exchanges should be strengthened in the future. The current research hotpots of SFTSV is antiviral therapy, immunotherapy, virus transmission mechanism and immune response.

KEYWORDS

SFTSV, bibliometric, data visualization, CiteSpace, VOSviewer

1. Introduction

Severe fever with thrombocytopenia syndrome (SFTS) virus (SFTSV) is a type of Bunyavirus that is seemingly transmitted by ticks, such as *Haemaphystick longicornis* (1, 2). A significant number of SFTSV are human pathogens that can cause severe diseases such as hepatitis, encephalitis, and hemorrhagic fever (3). SFTSV was firstly detected in 2009 in rural parts of

central and northeastern China (4). It has an incubation period of 5–14 days and patients present with clinical symptoms such as fever (>38.6°C), thrombocytopenia, leukopenia, and gastrointestinal symptoms (5). Some severe cases may also present with disturbance of consciousness, skin petechiae, and gastrointestinal and pulmonary hemorrhage (6). In severe cases, secondary encephalopathy and multiple organ failure, fulminant myocarditis, rhabdomyolysis, and hemophagocytic syndrome may develop. This emerging infectious zoonotic disease has also been reported in other Asian countries such as South Korea and Japan (7). In addition, a report in 2012 showed that the clinical manifestations of two cases of new venous virus infection in Missouri, USA were similar to that of SFTSV, and phylogenetic analysis also showed that the virus found in the USA was closely related to SFTSV (8).

The fatality rate among the first infected patients in China was reported to be 30% (7). In response to this emerging infectious disease of unknown cause, in 2010, enhanced surveillance for severe fever with SFTS was implemented. In 2015, the fatality rate for SFTS in Japan and South Korea was over 30% (9). SFTSV can evolve rapidly through genetic mutations and has already become a major threat to public health (3). There is no specific treatment for SFTSV, and the only method to prevent SFTSV infection and transmission is to avoid tick bites (10). Moreover, there is currently no vaccine for SFTSV. Therefore, we must understand the related risk factors and outcomes to strengthen our preparedness strategies.

We performed bibliometric analysis of SFTSV publications published in the Web of Science Core Collection (WoSCC) and

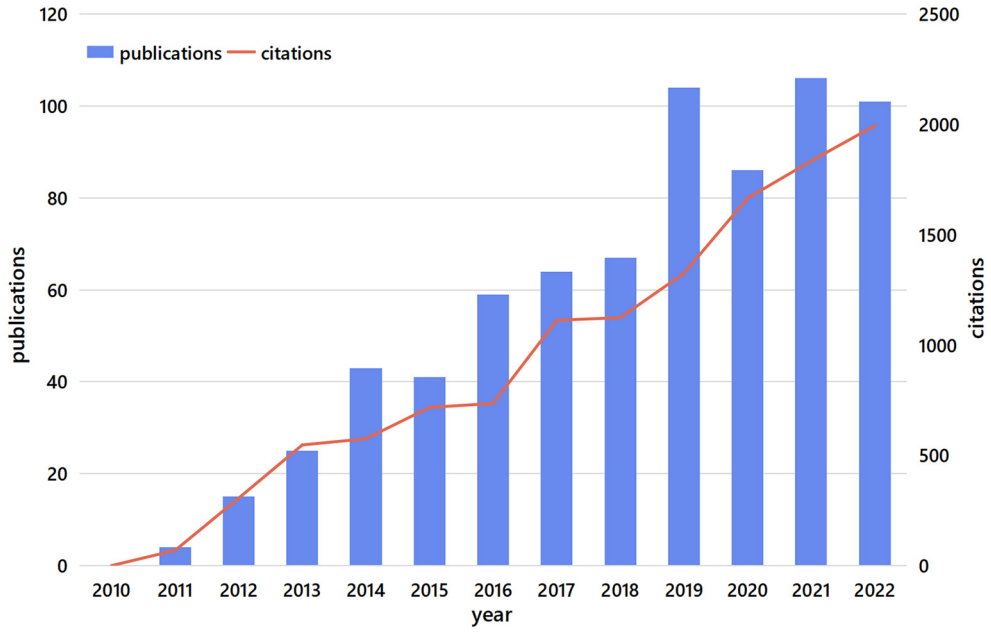


FIGURE 1
Trend of publications and citations (2011–2022).

TABLE 1 The top 10 countries in number of publications concerning SFTSV.

Rank	Country	Total publications	Total citations	Average citations	TLS
1st	Peoples R China	359	8,085	22.5	105
2nd	Japan	162	2,216	13.7	41
3rd	South Korea	143	1,852	13.0	35
4th	the United States	104	3,129	30.1	101
5th	Germany	17	346	20.4	14
6th	United Kingdom	12	392	32.7	15
7th	France	9	130	14.4	11
8th	India	9	68	7.6	7
9th	Zambia	5	101	20.2	8
10th	Australia	4	11	2.8	6

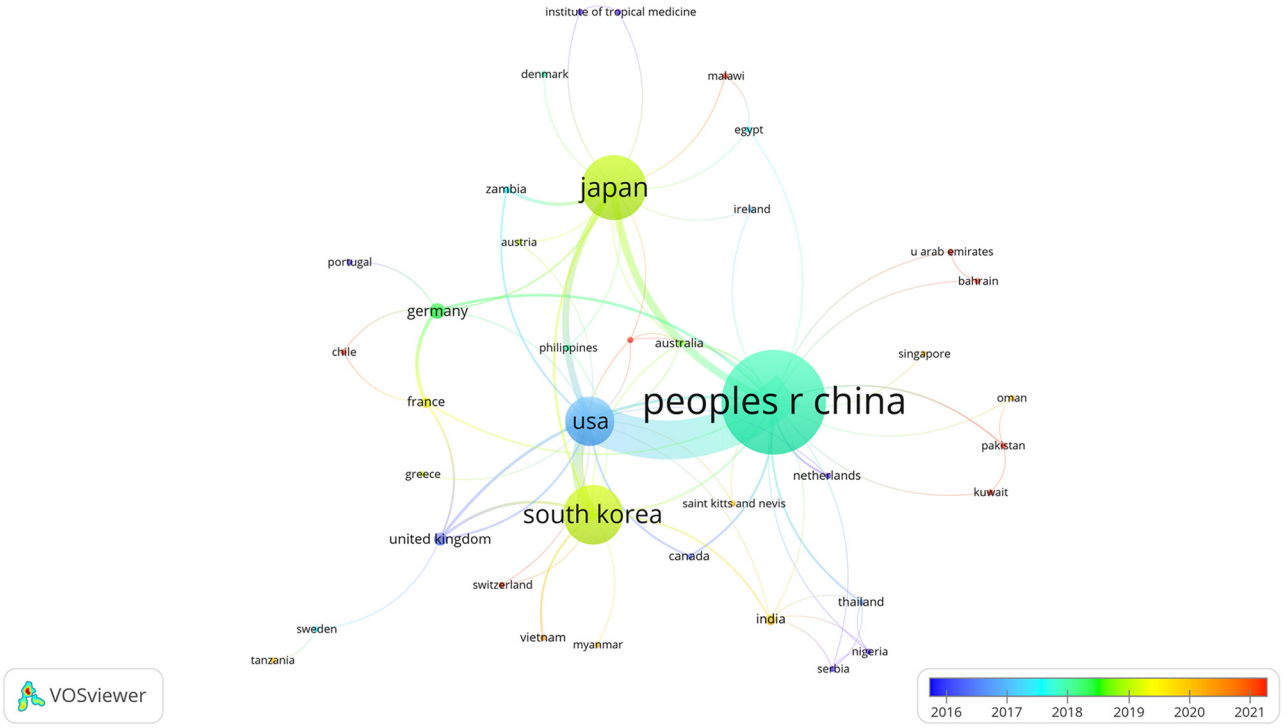


FIGURE 2
Cooperation map of countries.

TABLE 2 The top 10 journals in number of publications and citations concerning SFTSV.

Rank	Publication journal	Documents	IF*	Citations	Cited journal	Documents	IF*	Co-Citations
1st	Viruses-Basel	34	5.818	297	Journal of Virology	29	6.549	1,586
2nd	Emerging Infectious Diseases	31	16.126	1,332	Emerging Infectious Diseases	31	16.126	1,352
3rd	Journal of Virology	29	6.549	1,199	Journal of Infectious Diseases	6	7.759	807
4th	PLoS Neglected Tropical Diseases	28	4.781	369	New England Journal of Medicine	1	176.077	772
5th	PLoS ONE	24	3.752	498	Clinical Infectious Diseases	5	20.999	685
6th	Ticks and Tick-Borne Diseases	23	3.817	377	PLoS ONE	24	3.752	579
7th	American Journal of Tropical Medicine and Hygiene	19	3.707	375	Proceedings of the National Academy of Sciences of the United States of America	4	12.778	411
8th	Scientific Reports	17	4.996	233	American Journal of Tropical Medicine and Hygiene	19	3.707	411
9th	Frontiers in Microbiology	17	6.064	111	PLoS Neglected Tropical Diseases	28	4.781	388
10th	Japanese Journal of Infectious Diseases	16	2.541	188	PLoS Pathogens	9	7.464	329

*Impact factors (IF) based on Clarivate Analytics 'Journal Citation Reports (JCR) 2021.

Scopus databases. We assessed the number and year of publications, the number of national or regional contributions, the number of journals and disciplines, the number of institutions, the number of international collaborations, and the number of authors, and performed a reference analysis. By analyzing the co-occurrence of SFTSV keywords on a visualization diagram, the SFTSV-related research trends and hotspots were determined, and we hope to encourage advancements in the field through this work. In this report, we review and describe the development and main trends in the published SFTSV literature with a view to providing a new perspective and reference for the future of SFTSV research and SFTSV prevention.

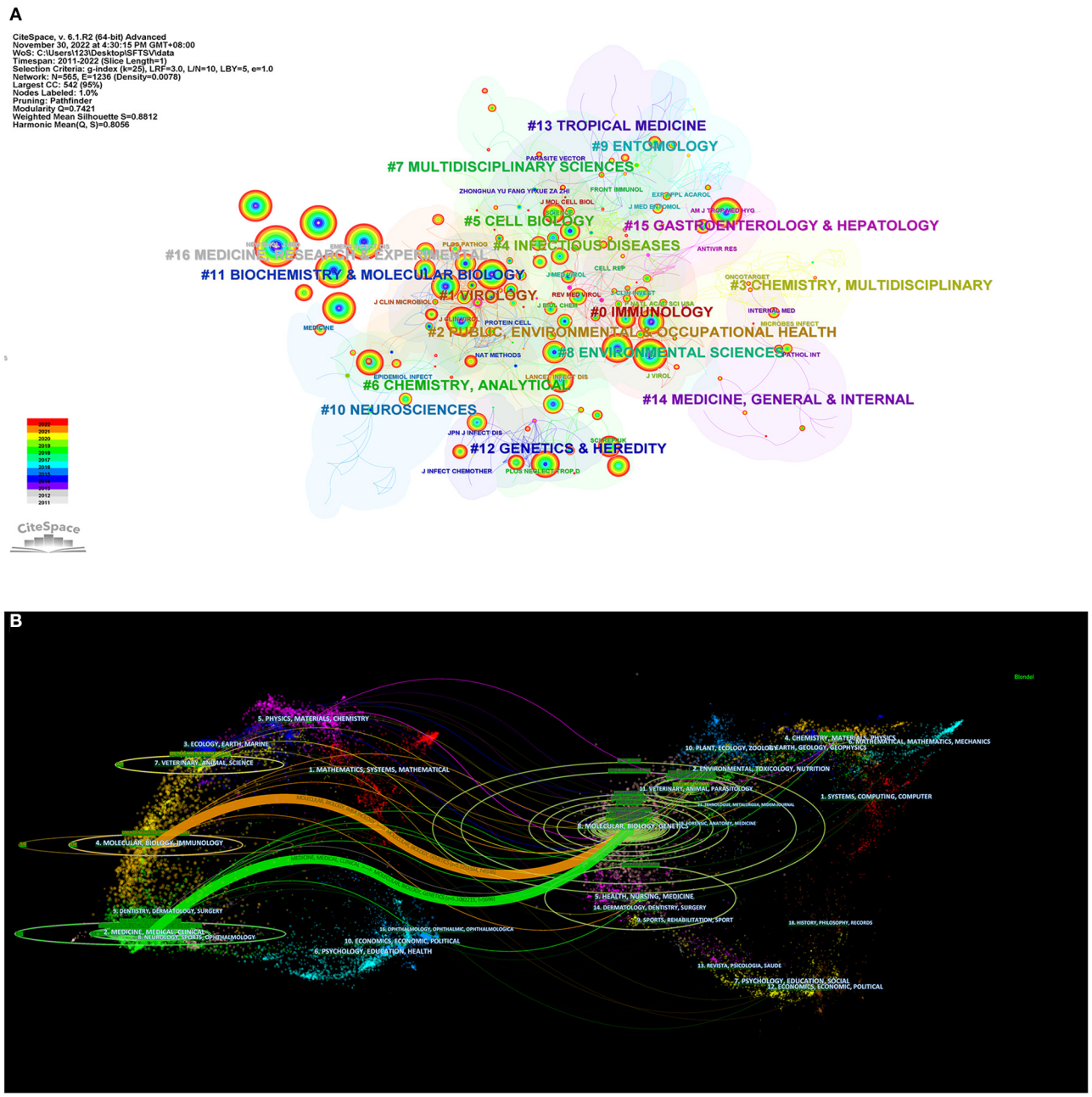


FIGURE 3
Contributions of top journals on SFTSV. **(A)** Journal co-citation network of clustering subjects categories. **(B)** A dual-map overlap of journals on SFTSV research by citespace (distribution of cited literature and journals on the left and articles that cited literature on the right).

2. Materials and methods

2.1. Data sources and search strategy

The search terms used were "severe fever with thrombocytopenia syndrome bunyavirus" or "severe fever with thrombocytopenia syndrome virus," or "SFTS virus," or "SFTSV," and the article titles, abstract, and keywords of publications retrieved from the WoSCC and Scopus databases were searched (11, 12). The search period was form January 1, 2010 to November 21, 2022. The document type was limited to "article" and "review article," and the language was "English." The three researchers simultaneously examined the articles concerning STFSV by title, abstract, and keywords, and read the full paper if necessary. A total of 648 and 707 publications were identified from the WoSCC and Scopus databases, respectively (2011–2022), see Supplementary Figure 1 for details. The data used in this study came from a public database, so the approval of the ethical committee was not required.

2.2. Data analysis and visualization

To ensure the consistency of the results and the reproducibility of this study, two researchers analyzed the data separately. Microsoft Excel 2020 was used to analyze and represent the most productive

TABLE 3 The top 10 institution in number of publications concerning SFTSV.

Rank	Institution	Countries/regions	Documents	Citations	Centrality
1st	National Institute of Infectious Diseases	Japan	97	1,251	0.25
2nd	Chinese Academy of Sciences	Peoples R China	58	1,434	0.35
3rd	Chinese Center for Disease Control and Prevention	Peoples R China	53	2,108	0.41
4th	Beijing Institute of Microbiology and Epidemiology	Peoples R China	47	1,033	0.21
5th	Shandong University	Peoples R China	32	1,032	0.07
6th	Jiangsu Provincial Center for Disease Control and Prevention	Peoples R China	32	1,037	0.13
7th	Anhui Medical University	Peoples R China	32	525	0.04
8th	University of Texas Medical Branch	the United States	30	1,579	0.15
9th	Peking University	Peoples R China	29	426	0.05
10th	Seoul National University	South Korea	29	656	0.15

and cited authors, institutions, journals, and countries/regions. All the data from the WoSCC and Scopus databases were processed using the following functionality of the CiteSpace software: convert to Web of Science plain format and remove duplicates through the dol.list file of the citing article. CiteSpace was used to analyze and display the development context and research hotspots in the STFSV field, and to predict its evolutionary paths, as well as research frontiers (13, 14). The citation bursts for publication year, author, institution, journal, country/region, and keywords were identified and visualized using CiteSpace. The knowledge graph of the parameter "years per slice" was adjusted to 1 year. VOSviewer was used to draw visual knowledge maps for distributions of authors, institutions, and countries/regions, as well as disciplines (15).

In the knowledge graph, different nodes represented various elements such as authors, institutions, countries/regions, and keywords. The size of the nodes reflected the number or frequency of publications, and the larger the node, the higher the number or frequency of publications. The connection lines between the nodes reflected the relationship between the co-operation or co-citation, the thicker the line, the more times of cooperation or co-citation. The different colors of the lines within the nodes represented different times, and the color of the line indicated the years when the co-operation or co-citation first appeared (16–18).

3. Results

3.1. Trends of publications and citations

The search results were identified by source and duplicates were removed. A total of 715 unique records were obtained, including 672 articles (94.0%) and 43 reviews (6.0%), from WoSCC (594) and Scopus (121). Since 2011, the number of SFTSV-related publications increased rapidly and attained a maximum peak in 2021. The number of citations showed a similar but more stable upward trend compared to the number of publications. The publication in 2022 is less than that in 2021, because this study started in November 2022 and has not finished in 2022, so it is slightly less than that in 2021 (Figure 1).

3.2. Contributions of countries/regions

A total of 45 countries/regions have published relevant publications. The Peoples Republic of China ranked first with 359 documents (50.2%), followed by Japan (162, 22.7%) and South Korea (143, 20.0%). The highest total citations of publications were published from The Peoples Republic of China, but the average citations per article was lower than the United States and United Kingdom (Table 1). Total link strength (TLS) indicated the number of connections between nodes. Stronger TLS means more cooperation with other countries. As is shown in Table 1, Peoples Republic of China and the United States were the two countries that carry out the most international cooperation.

When the minimum number of documents for VOSviewer was set to 1, 39 countries/regions met the thresholds. Superimposition of time on the visualized cooperation map of countries/regions is shown in Figure 2. The more purple the color, the earlier the country appeared, and the redder the color, the later the country appeared. The thicker the line is, the stronger the cooperation. The top four total link strength countries/regions were The Peoples Republic of China, the United States, Japan, and South Korea. The Peoples Republic of China had cooperated with numerous countries/regions in SFTSV-related research, and their most significant cooperations were with the United States and Japan.

3.3. Contributions of top journals

All the publications were published in a total of 207 journals, and 34 journal published at least five articles, with the top 10 journals accounting for 33.3% of the total publications included in this analysis. Among the top 10 publication journals on SFTSV, Viruses-Basel (IF = 5.818) ranked highest with ~34 articles, followed by Emerging Infectious Diseases (IF = 16.126, 31), and Journal of Virology (IF = 6.549, 29). Journal of Virology (IF = 6.549), Emerging Infectious Diseases (IF = 16.126, 31), and Journal of Infectious Diseases (IF = 7.759) were the top 3 ranked cited journals with 1,586, 1,352, and 807 co-citations, respectively (Table 2).

The journals that published SFTSV-related articles were mainly from the immunology, virology, public, environmental and occupational health fields. The different colors of the circles indicate

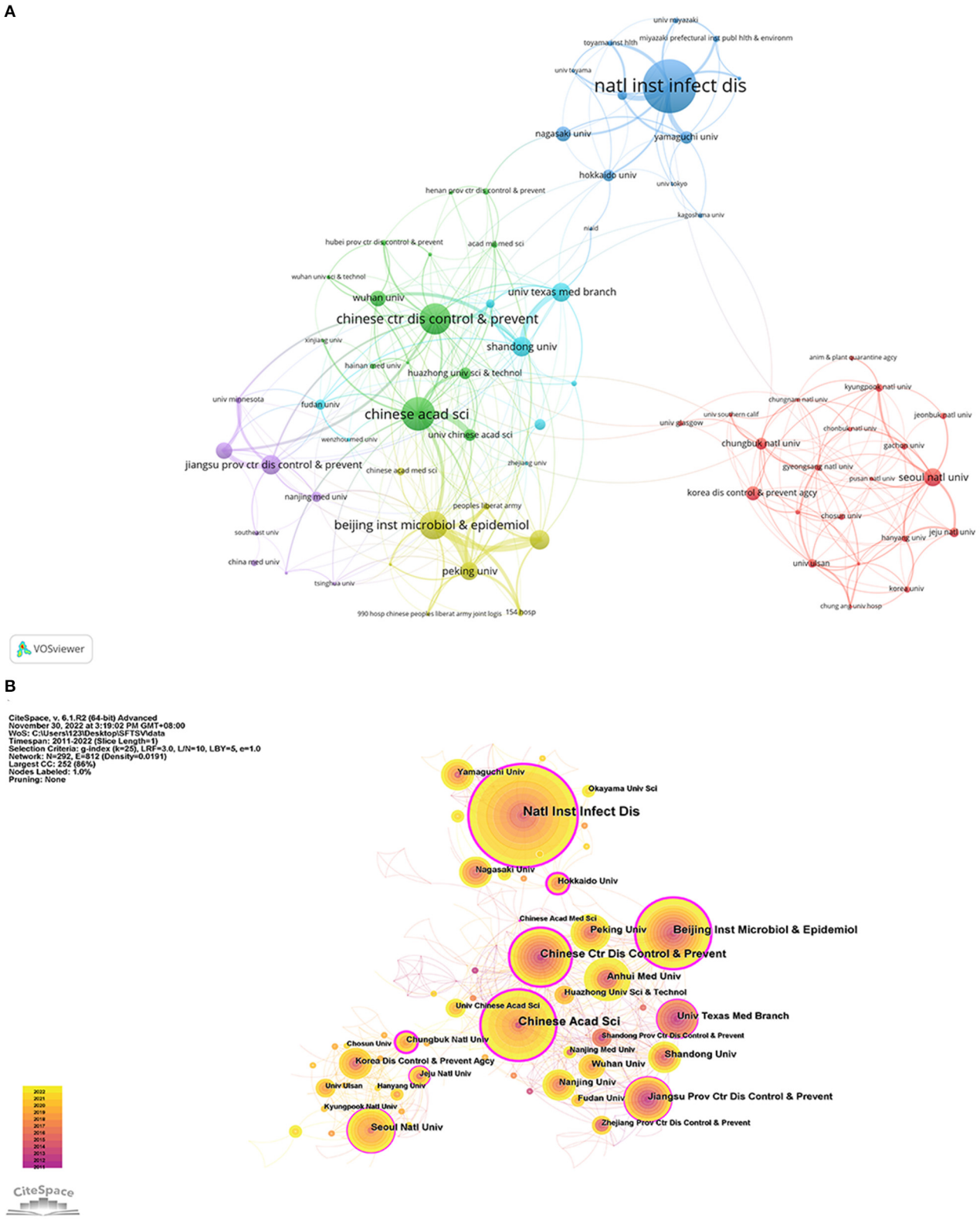


FIGURE 4
Visualization map of institution on SFTSV. **(A)** Cooperation map of 72 institutions with the number of publications no less than 5 times. **(B)** Centrality cooperation map of institutions.

the number of citations of the journals in different years, with purple to red representing the year of the citation from oldest to most recent (Figure 3A).

The dual map of the journals shows that there were two paths for citing and cited journals: (1) Molecular, Biology and Immunology-Molecular, Biology, Genetics; and (2) Medicine,

TABLE 4 The top 10 authors in number of publications concerning SFTSV.

Rank	Author	Countries/ regions	Documents	Citations
1st	Saijo Masayuki	Japan	43	829
2nd	Shimojima Masyuki	Japan	40	703
3rd	Liu Wei	Peoples R China	36	861
4th	Deng Fei	Peoples R China	29	555
5th	Liang Mifang	Peoples R China	28	1,453
6th	Yu Xuejie	Peoples R China	28	1,163
7th	Morikawa Sshigeru	Japan	27	519
8th	Cui Ning	Peoples R China	26	80
9th	Li Dexin	Peoples R China	25	1,339
10th	Li Hao	Peoples R China	24	522

Medical, Clinical-Molecular, Biology, Genetics. The citing journals were mainly concentrated in four circles including six fields: (1) Chemistry, Materials, Physics; (2) Veterinary, Animal, Parasitology; (3) Molecular, Biology, Genetics; (4) Health, Nursing, Medicine; (5) Dermatology, Dentistry, Surgery; and (6) Sports, Rehabilitation, Sport (Figure 3B).

3.4. Analysis of institution and co-institution

The "institution" node analysis showed that the institution with the largest number of SFTSV-related articles published was the National Institute of Infectious Disease (Japan), and seven of the top 10 research institutions were located in The Peoples Republic of China (Table 3). The VOSviewer parameters were set to the minimum number of documents for an institution = 5, and 72 institutions were obtained. As shown in Figure 4A, different colors represent different collaboration clusters, almost all of which were conducted within individual countries, with South Korea in red, Japan in blue, and Peoples Republic of China in green, yellow and purple.

CiteSpace was used to construct the co-institutions knowledge map (Figure 4B). The purple outermost circle represents the betweenness centrality (BC) (19), which indicates the significance of the nodes in networks. The larger the purple circle, the greater the BC, which is an indicative measure of the contribution of research achievements of institutions to the SFTSV field. Chinese Academy of Sciences (BC = 0.41), Chinese Center for Disease Control and Prevention (BC = 0.35), and Chungbuk National University (BC = 0.32) occupied important positions in the cooperation network.

3.5. Analysis of author

The top three SFTSV-related researchers with the largest number of published articles were Saijo, Shimojima, and Liu Wei, with 43, 40, and 36 articles respectively (Table 4). SFTSV-related researchers with more than 10 published articles were included in the network of co-authorship. Cooperations among authors were divided into clusters with different colors based on co-authorship analysis. As shown in Figure 5, the 55 authors were divided into 10 clusters with different colors. The connection lines between the different clusters were very thin or had no connection lines, indicating little or no cooperation between the clusters.

3.6. Analysis of reference

Using cited reference analysis of VOSviewer, the minimum number of documents of a cited reference was set at 20, and 148 met the threshold. The cited reference network with four clusters is shown in Figure 6A. McMullan et al. (8), Kim (20), Yu et al. (21), and Takahashi et al. (22) had the highest total citations in the red, yellow, green, and blue clusters respectively. Yu et al.'s (21) article had the largest number of co-citations, indicating that this article is the most important research achievement in the SFTSV field (Table 5).

A co-cited references cluster analysis was conducted using CiteSpace with set node type = cited reference, and other parameters were set to their default values. We obtained 22 clusters that clearly demonstrated the research themes in the SFTSV field (Figure 6B), and different colors from gray to red represented the number of co-citations in different years. Figure 6C presents the top 20 references with the strongest citation bursts. "Fever with thrombocytopenia associated with a novel bunyavirus in China," which began in 2011, was the strongest citation burst with an intensity of 58.19.

3.7. Analysis of keywords

Keyword co-occurrence analysis was used to determine the main directions and hotspots in SFTSV-related research. In the co-occurrence analysis in VOSviewer, more than 121 keywords occurred more than 20 times. The keyword co-occurrence network map of three different color clusters shows the main directions in SFTSV research. The keywords in the red cluster were mainly pathogeneses and symptoms, the green cluster was related to the characteristics of the virus and infected patients, and the blue cluster was the transmission mechanism and high-risk factors in SFTSV (Figure 7A).

CiteSpace performed cluster calculation according to the co-occurrence of keywords, and obtained 11 clusters, as listed in Supplementary Table 1. The modularity Q = 0.5855 and mean silhouette S = 0.8464, which was considered high, meaning that the network was reasonably divided into loosely coupled clusters (Figure 7B). The keywords of 11 clusters were described along the horizontal timeline in the timeline visualization, showing the research progress in the SFTSV field from 2011 to 2022 (Figure 7C). New keywords appeared in different clusters from 2021 to 2022; for example: Arenaviridae, virus transmission, and Crimean Congo hemorrhagic fever virus appear in #5 virus hemorrhagic fever; creatinine and platelet count in #3 bunyavirus; adaptive immune response and transcription in #7 protein, which may indicate new direction in SFTSV research. Citation bursts are terms that occur abruptly or increased dramatically in frequency in a short period of time, indicating the evolution of the research hotspot over time. The citation bursts in this discipline began in 2011, and 25 keywords with the strongest citation bursts from 2011 to 2022. The strongest citation bursts were syndrome bunyavirus with 7.66 strength from 2014 to

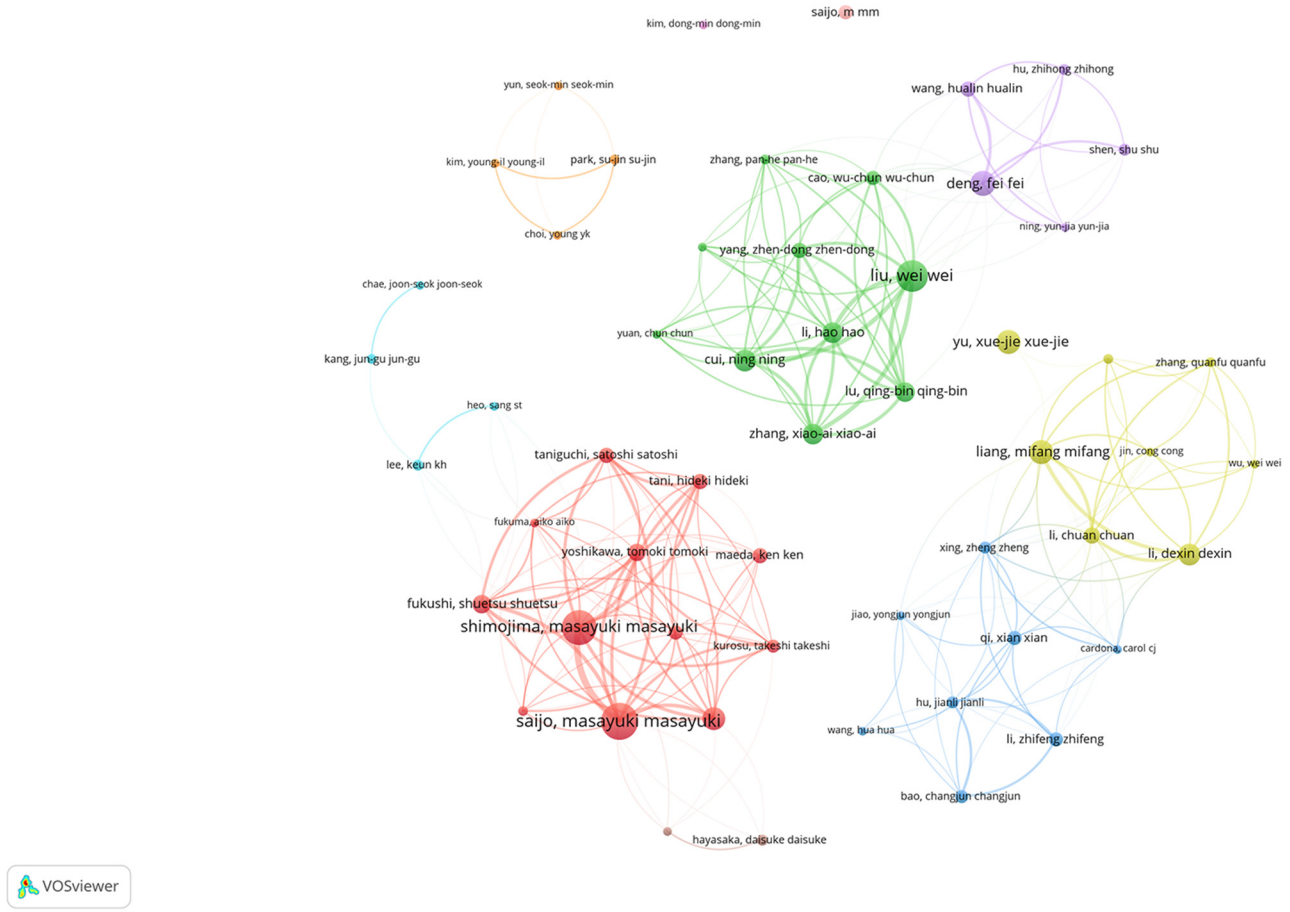


FIGURE 5
Cooperation map of 55 authors with the number of publications no less than 5 times.

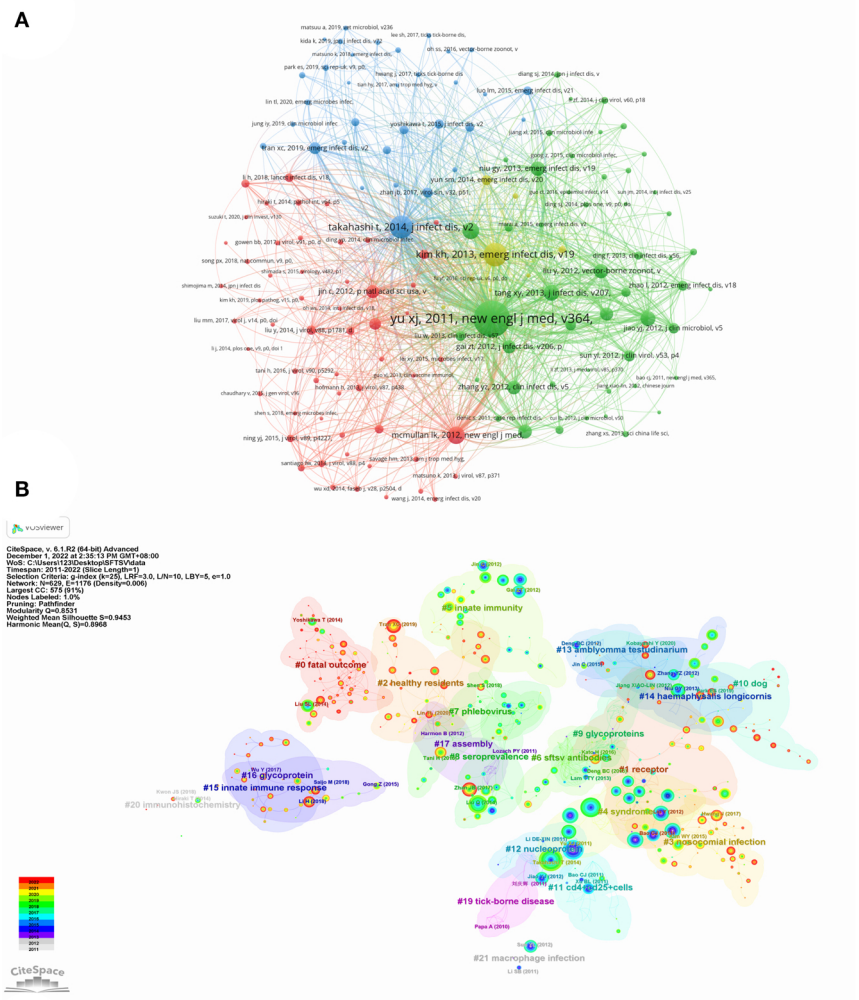
2016. Four keywords (response, cat, mechanism, and domesticated animal) appeared at the end of 2022, and most probably represent hotspots in current SFTSV research (Figure 7D).

4. Discussion

We found a total of 715 unique published articles on SFTSV research across two databases. The number of publications showed a significant upward trend since 2011. This trend may be related to the widespread epidemic of SFTSV in many countries/regions in Asia after 2010 (23). The first case of SFTSV was reported from The Peoples Republic of China in 2009, and further cases have been reported in many countries in Asia over the past 10 years, including South Korea and Japan (2013), Vietnam (2017), Myanmar (2018), Taiwan region (2019), and Thailand and Pakistan (2020) (21, 22, 24–29). Since the first case report, SFTSV has attracted great public health attention in Asia, due to the large number of case reports and its high initial case fatality rate of 12% to 30% (21, 30).

The Peoples Republic of China had the largest number of SFTSV cases reported and was the most widespread. A total of 13,824 SFTSV cases (8,899 laboratory-confirmed cases and 4,925 probable cases) had been reported in The Peoples Republic of China, and the regional distribution of SFTSV in The Peoples Republic of China had gradually expanded from five provinces in 2010 to 25 provinces in 2019 (30). The Peoples Republic of China was the main contributor and the most important cooperative partner in the field of SFTSV globally. The Peoples Republic of China had published the most articles in the SFTSV field in the past decade or more (359 as of November 2022). The Peoples Republic of China had cooperated with many countries in SFTSV research, as shown in Figure 2. The United States is the most important cooperative country for The Peoples Republic of China in SFTSV-related research, mainly related to the Heartland Virus, a virus similar to SFTSV isolated from two independent patients in northwest Mississippi in the United States in 2012 (8). In this study, SFTSV-related articles were mainly published in four countries, namely, The Peoples Republic of China, Japan, South Korea, and the United States. However, although Japan and South Korea ranked the third and fourth with 162 and 143 articles published, respectively, and the total citations and the average citations were much lower than those in the Peoples Republic of China and the United States, indicating that Japan and South Korea lack high-quality articles in SFTSV research.

The research findings of collaborations were also influenced by country. The Peoples Republic of China had seven of the top 10 institutions (Table 3). A number of universities, research laboratories, and even government departments located in the Peoples Republic of China had made contributions to SFTSV research, which



C

Top 25 References with the Strongest Citation Bursts

References	Year	Strength	Begin	End	2011 - 2022
Yu XJ, 2011, NEW ENGL J MED, V364, P1523, DOI 10.1056/NEJMoa1010095, DOI	2011	58.19	2011	2016	
Bao CJ, 2011, CLIN INFECT DIS, V53, P1208, DOI 10.1093/cid/cir732, DOI	2011	26.17	2012	2016	
Liu Y, 2012, VECTOR-BORNE ZOONOT, V12, P156, DOI 10.1089/vbz.2011.0758, DOI	2012	18.45	2012	2017	
Xu BL, 2011, PLOS PATHOG, V7, P0, DOI 10.1371/journal.ppat.1002369, DOI	2011	17.86	2012	2016	
Zhang YZ, 2012, CLIN INFECT DIS, V54, P527, DOI 10.1093/cid/cir804, DOI	2012	15.74	2012	2016	
Gai ZT, 2012, CLIN INFECT DIS, V54, P249, DOI 10.1093/cid/cir776, DOI	2012	15.27	2012	2015	
Jiao YJ, 2012, J CLIN MICROBIOL, V50, P372, DOI 10.1128/JCM.01319-11, DOI	2012	14.62	2012	2017	
Sun YL, 2012, J CLIN VIROL, V53, P48, DOI 10.1016/j.jcv.2011.09.031, DOI	2012	12.83	2012	2015	
Zhang YONG-ZHEN, 2011, ZHONGHUA LIU XING BING XUE ZA ZHI, V32, P209	2011	11.49	2012	2016	
Mcmullan LK, 2012, NEW ENGL J MED, V367, P834, DOI 10.1056/NEJMoa1203378, DOI	2012	22.76	2013	2017	
Zhao L, 2012, EMERG INFECT DIS, V18, P963, DOI 10.3201/eid1806.111345, DOI	2012	14.9	2013	2017	
Gai ZT, 2012, J INFECT DIS, V206, P1095, DOI 10.1093/infdis/jis472, DOI	2012	13.13	2013	2017	
Zhang YZ, 2012, J VIROL, V86, P2864, DOI 10.1128/JVI.06192-11, DOI	2012	12.22	2013	2017	
Jin C, 2012, P NATL ACAD SCI USA, V109, P10053, DOI 10.1073/pnas.1120246109, DOI	2012	11.15	2013	2017	
Kim KH, 2013, EMERG INFECT DIS, V19, P1892, DOI 10.3201/eid1911.130792, DOI	2013	26.31	2014	2018	
Tang XY, 2013, J INFECT DIS, V207, P736, DOI 10.1093/infdis/jis748, DOI	2013	15.16	2014	2017	
Niu GY, 2013, EMERG INFECT DIS, V19, P756, DOI 10.3201/eid1905.120245, DOI	2013	13	2014	2018	
Takahashi T, 2014, J INFECT DIS, V209, P816, DOI 10.1093/infdis/jit603, DOI	2014	17.45	2016	2019	
Liu Q, 2014, LANCET INFECT DIS, V14, P763, DOI 10.1016/S1473-3099(14)70718-2, DOI	2014	16.22	2016	2019	
Kato H, 2016, PLOS ONE, V11, P0, DOI 10.1371/journal.pone.0165207, DOI	2016	11.18	2018	2022	
Tran XC, 2019, EMERG INFECT DIS, V25, P1029, DOI 10.3201/eid2505.181463, DOI	2019	28.73	2020	2022	
Li H, 2018, LANCET INFECT DIS, V18, P1127, DOI 10.1016/S1473-3099(18)30293-7, DOI	2018	15.74	2020	2022	
Park ES, 2019, SCI REP-UK, V9, P0, DOI 10.1038/s41598-019-48317-8, DOI	2019	13.46	2020	2022	
Kida K, 2019, JPN J INFECT DIS, V72, P356, DOI 10.7883/yoken.JJID.2018.526, DOI	2019	11.67	2020	2022	
Matsuu A, 2019, VET MICROBIOL, V236, P0, DOI 10.1016/j.vetmic.2019.06.019, DOI	2019	10.96	2020	2022	

FIGURE 6
Visualization map of reference on SFTSV. **(A)** Distribution of 148 reference with a frequency of no less than 20 times. **(B)** Reference co-citation network for clustering title terms. **(C)** Top 25 references with the strongest citation bursts.

demonstrates that Chinese institutions attach great importance to SFTSV. The Chinese Academy of Sciences and the Chinese Center for Disease Control and Prevention, which are located in the Peoples Republic of China play the most important role in the SFTSV field. The collaborative network of institutions and authors is loose, and a large number of studies are conducted mainly within one country,

TABLE 5 The top 10 reference in number of co-citations concerning SFTSV.

Rank	Author	Journals	DOI	Year	Co-citation
1st	Yu XJ	New Engl J Med	10.1056/NEJMoa1010095	2011	502
2nd	Tim KH	Emerg Infect Dis	10.3201/eid1911.130792	2013	260
3rd	Takahashi T	J Infect Dis	10.1093/infdis/jit603	2014	250
4th	Mcmullan LK	New Engl J Med	10.1056/NEJMoa1203378	2012	167
5th	Liu Q	Lancet Infect Dis	10.1016/S1473-3099(14)70718-2	2014	159
6th	Tang XY	J Infect Dis	10.1093/infdis/jis748	2013	144
7th	Bao CJ	Clin Infect Dis	10.1093/cid/cir732	2011	135
8th	Liu Y	Vector-Borne Zoonot	10.1089/vbz.2011.0758	2012	131
9th	Niu GY	Emerg Infect Dis	10.3201/eid1905.120245	2013	126
10th	Xu BL	PLoS Pathog	10.1371/journal.ppat.1002369	2011	124

illustrating the lack of international collaboration in SFTSV research. As SFTSV continues to be reported and valued in countries around the world, we believe that the institutions and authors from Asian countries should cooperate more extensively and closely.

Journals with more citations and co-citation frequencies play an important role in the SFTSV field. However, not all high-yield journals have high numbers of citations. The journal with the most published articles was Viruses-Basel, but the total citations was only 297, indicating that articles from this journal were not the main choice for most SFTSV-related researchers. The citations and co-citations of the Journal of Virology and Emerging Infectious Diseases were among the top 3, indicating that these two journals have published high-quality publications with convincing results. Table 2 lists the top 10 high-yield journals and high co-citation journals for SFTSV-related research, and academic publications on SFTSV may be preferentially published in high-yield journals, while high co-citation journals had published mature research results. There are only two paths between the cited journal and the citing journal, which means that advancement in the SFTSV field will require more cross-disciplinary collaboration.

Yu's (21) article, "Fever with thrombocytopenia associated with a novel bunyavirus in China" published in the New England Journal of Medicine in 2011 had the highest number of co-citations. The article introduced the process of isolating the virus from the blood sample, identified the family of the virus based on the RNA sequence analysis, and named it SFTS bunyavirus. The presence of the virus was confirmed from 171 patients by detection of viral RNA or specific antibodies to the virus in the blood, or both. This was the first systematic article on the isolation and diagnosis of SFTSV and is the most important research achievement in the SFTSV field.

Keyword frequency may reflect the development tendency of research hotspots from another point of view. VOSviewer divided 715 previous studies into three categories according to keywords: (1) pathogenesis and symptoms, (2) characteristics of the virus and infected patients, and (3) transmission mechanism and risk factors of SFTSV.

The main clinical features of SFTS on presentation include fever, thrombocytopenia, leukocytopenia, and gastrointestinal symptoms (31). Before 2010, SFTSV was assumed to be caused by bacteria such as *Anaplasma phagocytophilum* bacterium, due to the fact that no pathogens had been isolated from patients (32). Virus isolation is the commonly employed method in research institutions and it provides sufficient evidence of SFTSV infection, but it is very time-consuming. Amplification of viral nucleic acid and the reverse transcription PCR method are frequently used for clinical confirmation (33).

According to the pathogenicity of the bunyavirus, SFTSV prevents the host's immune response and is manifested through intense virus replication along with multiple organ failure (34). Examination of SFTSV patients has revealed that the number of natural killer cells increases, mainly during the acute phase and grievous SFTSV infection (35). Natural killer cells play an immunomodulatory function by producing various cytokines, and the number of these cytokines is proportional to the severity of the disease. Natural killer cells perform their immunomodulatory functions by producing various cytokines, the amount of which is proportional to the severity of the disease. These cytokines play important roles in serum virus load and other related clinical characteristics. Monocyte chemotactic protein 1 and interleukin-8 are crucial in progressive kidney injury (36), monocyte chemotactic protein 1 and interferon-c-inducible protein produce liver inflammation with fibrosis (37), and interleukin-8 can raise capillary permeability (38).

Much of the SFTSV research focused on "to person transmission" SFTSV is an infectious phlebovirus with a high mortality rate. It mainly affects humans, and although SFTSV can be transmitted through infected animals, there have been no reported cases of SFTSV infection in animals. Vertebrates are the host of this disease and ticks function as vectors, where the virus can undergo brisk changes using gene mutation, homologous recombination, and reassortments. Although the mechanism of SFTSV transmission remains unclear, a large number of studies have suggested that ticks may be the transmission vectors and that domestic or wild animals may be amplifying hosts (39). Age was considered to be a critical risk factor for morbidity and mortality in SFTS, which mainly targets people over 50 years of age. Farmers were the main high-risk population, and there were more female cases than male cases (30).

Based on the timeline viewer, we combined the keywords cluster with the relevant article for analysis to find the recent hotpots of SFTSV-related research. Five clusters emerged with new keywords in 2022.

Antiviral therapy in #1 SFTSV. Zhang et al.'s review (40) suggested that broad-spectrum antivirals have the potential to

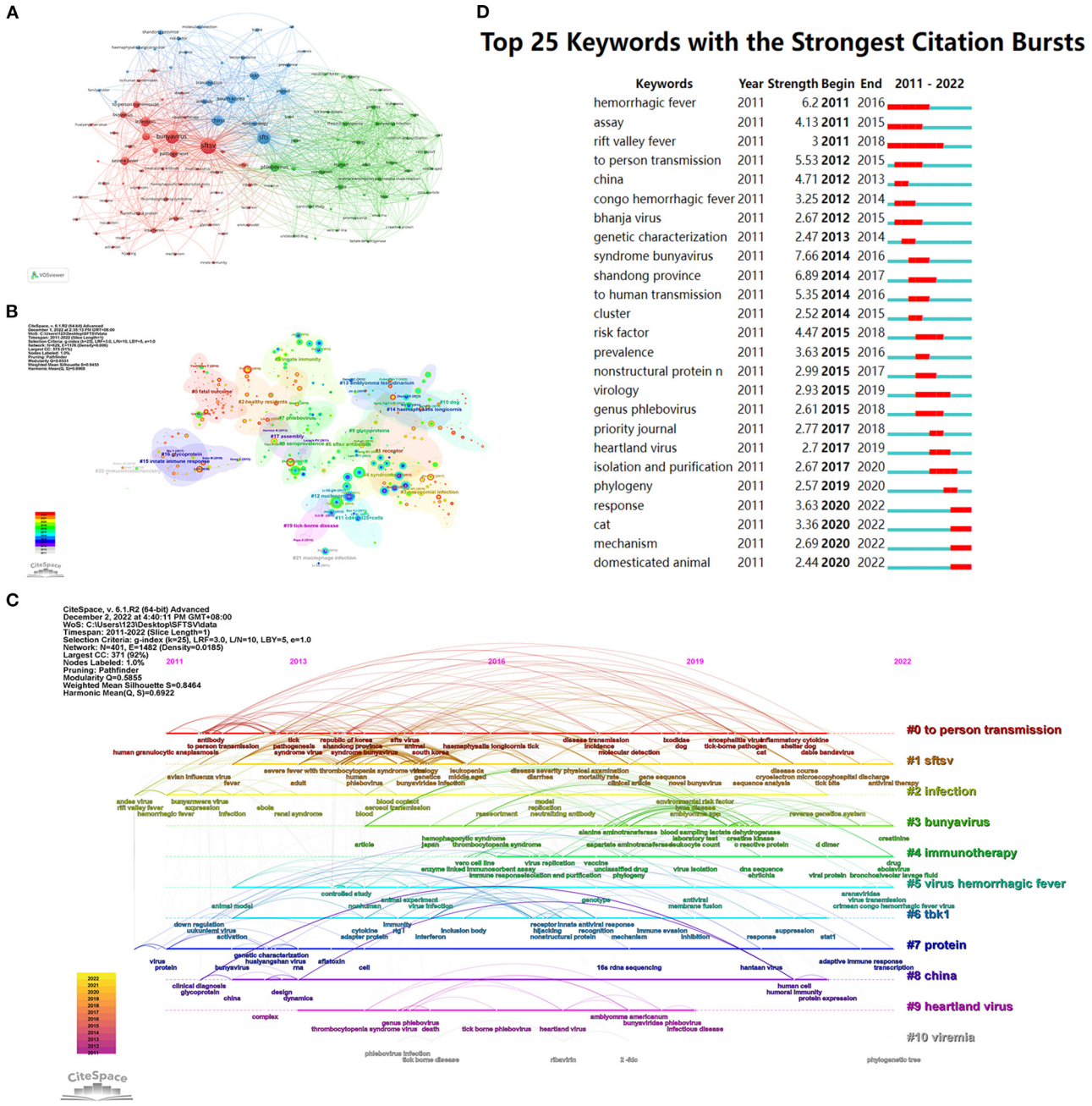


FIGURE 7
Visualization map of keywords on SFTSV. **(A)** Distribution of 121 keywords with a frequency of no less than 20 times. **(B)** Keywords co-occurrence network. **(C)** Timeline view of keyword cluster. **(D)** Top 25 keywords with the strongest citation bursts.

be the first line of defense to prevent the progression of the disease, and fapiravir maybe the most promising treatment. The future development of antiviral methods may depend mainly on targeted therapies such as monoclonal antibodies and prevention through vaccination.

Drug, ebolavirus, and bronchoalveolar lavage fluid in #4 immunotherapy. An approved drug (tilorone) was highly effective in the treatment of SFTSV infection and may have the potential to be a "universal vaccine" for antiviral infections (41). Małkowska et al. (42) reported that RIG-I-like receptors are promising in the treatment of viral hemorrhagic fevers. Bronchoalveolar lavage fluid is one of the therapeutic tools available to combat viral diseases (ebolavirus, SFTSV), and one reported case of aspergillus isolated in bronchoalveolar lavage fluid showed that invasive fungal disease may accompany the early clinical course of SFTSV infection (43, 44).

Crimean Congo hemorrhagic fever (CCHFV), virus transmission, and Arenaviridae in #5 virus hemorrhagic fever. Teng et al. (45) evaluated the environmental suitability and transmission risk of major Bunyavirales viruses in China by mapping the geographical distribution of all 89 Bunyavirales viruses reported in China from January 1951 to June 2021, and highlighted that Hantaviruses, *Dabie bandavirus*, and CCHFV had the severest

disease burden. Two viruses, CCHFV and Rift valley fever virus (RVFV), may occur in local area outbreaks in China. Xinjiang and southwestern Yunnan had the highest environmental suitability to CCHFV occurrence, and southern China had the highest environmental suitability to RVFV transmission all year round.

Adaptive immune response and transcription in #7 protein. Wang et al. (46) summarized the mechanism of SFTSV evasion of the host immune response and highlighted that SFTSV can escape from host immune responses *via* multiple strategies, such as interfering with the number and function of innate and adaptive immune cells, inhibiting the inhibiting interferon signaling pathway, regulating the NF-κB signaling, and autophagy. Moreover, a proposed strategy against the virus may be to regulate the host dysfunctional immune cells. Lan et al. (47) identified two clicks from the Lassa virus microgenome (MG) system, F1204 and F1781, that effectively inhibited authentic lymphocytic choriomeningitis virus (LCMV) and SFTSV infections.

Phylogenetic tree in #10 viremia. Xu et al. (32) constructed a phylogenetic tree based on the M segments of SFTSV, which was epidemic in the Jiaodong area of the Shandong Province, and found local endemic strains were mainly C2 and C3 isolates of SFTSV, and epidemic strains showed relatively stable heredity.

In terms of the top keywords with the strongest citation bursts, response, cat, mechanism, and domesticated animal appeared at the end of 2022 (Figure 7D). We believe that the anti-SFTSV immune responses, molecular mechanism, and viral transmission from animals may represent the future direction of SFTSV research.

Our bibliometrics have certain limitations inherents. First, CiteSpace and VOSviewer are used for bibliometric analysis in this paper, and only analyzed the main conclusions rather than full text, which cannot completely replace system search. Secondly, the data we retrieved were all from WoSCC and Scopus, which are considered to be the most commonly used databases for bibliometric analysis. Although the data from WoSCC and Scopus could be representative of numerous information to a certain extent, some documents excluded from these two databases should be taken into consideration, and the search strategy is designed to favor the matching accuracy of the search, and citation counts were probably underestimated. Finally, in this study, we mainly utilized a quantitative analysis approach, and paid little attention to qualitative research. As a result, certain critical points and details may be missed. Nevertheless, in this study, all the maps based on the retrieved data can intuitively present the hotspots, evolution process and development trend of SFTSV, which could providing many important reference values for newcomers in this field.

5. Conclusions

Through bibliometric analysis and data visualization, we elucidated on the research progress, research hotspots, and research frontier in the SFTSV field. We have identified the countries, journals, institutions, authors, and representative publications that play an important role in this field. The findings inform investigations on SFTSV and identify the potential partners for interested researchers. Our study found that SFTSV research is in a phase of explosive development, and a large number of publications have been published in the past decade. There is a lack of collaboration between countries and institutions, and international collaboration and exchanges should be strengthened in the future. The research direction of SFTSV is on the therapeutic strategy of SFTSV, the transmission mechanism of the virus, and the immune response. The current research hotpots of SFTSV is antiviral therapy, immunotherapy, virus transmission mechanism and immune response.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ZZ and JT designed the research. ZZ, WJ, HQ, and HZ collected and organized data. ZZ, YY, and LW analyzed the data. ZZ, JT, and XW drafted the manuscript. YY and XW contributed to the critical revision of the manuscript. All authors contributed to the manuscript and approved the submitted version.

Funding

This study was funded by Natural Science Foundation of Zhejiang Province (grant number LQ21H190004).

Acknowledgments

We would like to thank Editage (www.editage.com) for English language editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1120462/full#supplementary-material

References

1. Zhuang L, Sun Y, Cui X-M, Tang F, Hu J-G, Wang L-Y, et al. Transmission of severe fever with thrombocytopenia syndrome virus by *Haemaphysalis longicornis* ticks, China. *Emerg Infect Dis.* (2018) 24:868–71. doi: 10.3201/eid2405.151435

2. Yun S-M, Song BG, Choi W, Roh JY, Lee Y-J, Park WI, et al. First isolation of severe fever with thrombocytopenia syndrome virus from *Haemaphysalis longicornis* ticks collected in severe fever with thrombocytopenia syndrome outbreak areas in the Republic of Korea. *Vector Borne Zoonotic Dis.* (2016) 16:66–70. doi: 10.1089/vbz.2015.1832

3. Wang M, Tan W, Li J, Fang L, Yue M. The endless wars: severe fever with thrombocytopenia syndrome virus, host immune and genetic factors. *Front Cell Infect Microbiol.* (2022) 12:808098. doi: 10.3389/fcimb.2022.808098

4. Duan Y, Wu W, Zhao Q, Liu S, Liu H, Huang M, et al. Enzyme-antibody-modified gold nanoparticle probes for the ultrasensitive detection of nucleocapsid protein in SFTSV. *Int J Environ Res Public Health.* (2020) 17:4427. doi: 10.3390/ijerph17124427

5. Li D. A highly pathogenic new bunyavirus emerged in China. *Emerg Microbes Infect.* (2013) 2:e1. doi: 10.1038/emi.2013.1

6. Winkelhorst D, Kamphuis MM, Kloet LCd, Zwaginga JJ, Oepkes D, Lopriore E. Severe bleeding complications other than intracranial hemorrhage in neonatal alloimmune thrombocytopenia: a case series and review of the literature. *Transfusion.* (2016) 56:1230–5. doi: 10.1111/trf.13550

7. Dualis H, Zefong AC, Joo LK, Singh NKD, Rahim SSSA, Avoi R, et al. Factors and outcomes in Severe Fever with Thrombocytopenia Syndrome (SFTS): A systematic review. *Ann Med Surg.* (2021) 67:102501. doi: 10.1016/j.amsu.2021.102501

8. McMullan LK, Folk SM, Kelly AJ, MacNeil A, Goldsmith CS, Metcalfe MG, et al. A new phlebovirus associated with severe febrile illness in Missouri. *N Engl J Med.* (2012) 367:834–41. doi: 10.1056/NEJMoa1203378

9. Dong JL, Joon-Ho H, Hail P, Ki YP. Measuring the natural rate of interest with financial gaps: the cases of Japan and South Korea, Japan. *World Econ.* (2020) 54:101009. doi: 10.1016/j.japwor.2020.101009

10. Tong Y, Wang Q, Fu Y, Li S, Zhang Z, Zhang Z, et al. Molecular identification of severe fever with thrombocytopenia syndrome viruses from tick and bitten patient in Southeast China. *Virol J.* (2020) 17:122. doi: 10.1186/s12985-020-01391-1

11. AlRyalat SAS, Malkawi LW, Momani SM. Comparing bibliometric analysis using pubmed, scopus, and web of science databases. *J Visual Exp.* (2019) 152:58494. doi: 10.3791/58494

12. Philippe M, Adèle P-H. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics.* (2016) 106:213–28. doi: 10.1007/s11192-015-1765-5

13. Synnestvedt MB, Chen C, Holmes JH. CiteSpace II: visualization and knowledge discovery in bibliographic databases. In: *AMIA Annual Symposium Proceedings AMIA Symposium.* (2005). p. 724–8.

14. Wei J, Liang G, Alex J, Zhang T, Ma C. Research progress of energy utilization of agricultural waste in China: bibliometric analysis by citespace. *Sustainability.* (2020) 12:812. doi: 10.3390/su12030812

15. Eck NJv, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics.* (2010) 84:523–38. doi: 10.1007/s11192-009-0146-3

16. Chen C. Science mapping: a systematic review of the literature. *J Data Inform Sci.* (2017) 2:1–40. doi: 10.1515/jdis-2017-0006

17. Jesús L, Adrián S, AntonioJosé M, MaríaElena P. Robotics in education: a scientific mapping of the literature in Web of Science. *Electronics.* (2021) 10:291. doi: 10.3390/electronics10030291

18. Zhou Q, Kong H-B, He B-M, Zhou S-Y. Bibliometric analysis of bronchopulmonary dysplasia in extremely premature infants in the web of science database using CiteSpace Software. *Front Pediatr.* (2021) 9:705033. doi: 10.3389/fped.2021.705033

19. Wu H, Cheng K, Tong L, Wang Y, Yang W, Sun Z. Knowledge structure and emerging trends on osteonecrosis of the femoral head: a bibliometric and visualized study. *J Orthop Surg Res.* (2022) 17:194. doi: 10.1186/s13018-022-03068-7

20. Kim KH, Yi J, Kim G, Choi SJ, Jun KI, Kim NH et al. Severe fever with thrombocytopenia syndrome, South Korea, 2012. *Emerg Infect Dis.* (2013) 19:1892–4. doi: 10.3201/eid1911.130792

21. Yu X-J, Liang M-F, Zhang S-Y, Liu Y, Li J-D, Sun Y-L, et al. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N Engl J Med.* (2011) 364:1523–32. doi: 10.1056/NEJMoa1010095

22. Takahashi T, Maeda K, Suzuki T, Ishido A, Shigeoka T, Tominaga T, et al. The first identification and retrospective study of severe fever with thrombocytopenia syndrome in Japan. *J Infect Dis.* (2014) 209:816–27. doi: 10.1093/infdis/jit603

23. Liu S, Chai C, Wang C, Amer S, Lv H, He H, et al. Systematic review of severe fever with thrombocytopenia syndrome: virology, epidemiology, and clinical characteristics. *Rev Med Virol.* (2014) 24:90–102. doi: 10.1002/rmv.1776

24. Yun S-M, Lee W-G, Ryou J, Yang S-C, Park S-W, Roh JY, et al. Severe fever with thrombocytopenia syndrome virus in ticks collected from humans, South Korea, 2013. *Emerg Infect Dis.* (2014) 20:1358–61. doi: 10.3201/eid2008.131857

25. Tran XC, Yun Y, An LV, Kim S-H, Thao NTP, Man PKC, et al. Endemic severe fever with thrombocytopenia syndrome, Vietnam. *Emerg Infect Dis.* (2019) 25:1029–31. doi: 10.3201/eid2505.181463

26. Win AM, Nguyen YTH, Kim Y, Ha N-Y, Kang J-G, Kim H, et al. Genotypic heterogeneity of *Orientia tsutsugamushi* in scrub typhus patients and thrombocytopenia syndrome co-infection, Myanmar. *Emerg Infect Dis.* (2020) 26:1878–81. doi: 10.3201/eid2608.200135

27. Peng S-H, Yang S-L, Tang S-E, Wang T-C, Hsu T-C, Su C-L, et al. Human Case of severe fever with thrombocytopenia syndrome virus infection, Taiwan, 2019. *Emerg Infect Dis.* (2020) 26:1612–4. doi: 10.3201/eid2607.200104

28. Kim UJ, Kim D-M, Kim SE, Kang SJ, Jang H-C, Park K-H, et al. Case report: detection of the identical virus in a patient presenting with severe fever with thrombocytopenia syndrome encephalopathy and the tick that bit her. *BMC Infect Dis.* (2018) 18:181. doi: 10.1186/s12879-018-3092-y

29. Zohaib A, Zhang J, Saqib M, Athar MA, Hussain MH, Chen J, et al. Serologic evidence of severe fever with thrombocytopenia syndrome virus and related viruses in Pakistan. *Emerg Infect Dis.* (2020) 26:1513–6. doi: 10.3201/eid2607.190611

30. Huang X, Li J, Li A, Wang S, Li D. Epidemiological characteristics of severe fever with thrombocytopenia syndrome from 2010 to 2019 in mainland china. *Int J Environ Res Public Health.* (2021) 18:3092. doi: 10.3390/ijerph18063092

31. Yang T, Huang H, Jiang L, Li J. Overview of the immunological mechanism underlying severe fever with thrombocytopenia syndrome. *Int J Mol Med.* (2022) 50:118. doi: 10.3892/ijmm.2022.5174

32. Xu J, Liu Y, Zhang F, Wang X, Huang W, Wu Y, et al. Analysis of cross neutralizing activity of antibodies from sera of severe fever with thrombocytopenia syndrome patients to deal with different genotype strains. *Front Microbiol.* (2022) 13:1020545. doi: 10.3389/fmicb.2022.1020545

33. Sun Y, Liang M, Qu J, Jin C, Zhang Q, Li J, et al. Early diagnosis of novel SFTS bunyavirus infection by quantitative real-time RT-PCR assay. *J Clin Virol.* (2012) 53:48–53. doi: 10.1016/j.jcv.2011.09.031

34. Matsuno K, Weisend C, Kajihara M, Matysiak C, Williamson BN, Simuunza M, et al. Comprehensive molecular detection of tick-borne phleboviruses leads to the retrospective identification of taxonomically unassigned bunyaviruses and the discovery of a novel member of the genus phlebovirus. *J Virol.* (2015) 89:594–604. doi: 10.1128/JVI.02704-14

35. Sun L, Hu Y, Niyonsaba A, Tong Q, Lu L, Li H, et al. Detection and evaluation of immunofunction of patients with severe fever with thrombocytopenia syndrome. *Clin Exp Med.* (2014) 14:389–95. doi: 10.1007/s10238-013-0259-0

36. Arazi A, Rao DA, Berthier CC, Davidson A, Liu Y, Hoover PJ, et al. The immune cell landscape in kidneys of patients with lupus nephritis. *Nat Immunol.* (2019) 20:902–14. doi: 10.1038/s41590-019-0398-x

37. Marra F, Tacke F. Roles for chemokines in liver disease. *Gastroenterology.* (2014) 147:577–94. doi: 10.1053/j.gastro.2014.06.043

38. Petreaca ML, Yao M, Liu Y, Defea K, Martins-Green M. Transactivation of vascular endothelial growth factor receptor-2 by interleukin-8 (IL-8/CXCL8) is required for IL-8/CXCL8-induced endothelial permeability. *Mol Biol Cell.* (2007) 18:5014–23. doi: 10.1091/mbc.e07-01-0004

39. Wang S, Li J, Niu G, Wang X, Ding S, Jiang X, et al. SFTS virus in ticks in an endemic area of China. *Am J Trop Med Hyg.* (2015) 92:684–9. doi: 10.4269/ajtmh.14-0008

40. Zhang Y, Huang Y, Xu Y. Antiviral treatment options for severe fever with thrombocytopenia syndrome infections. *Infect Dis Ther.* (2022) 11:1805–19. doi: 10.1007/s40121-022-00693-x

41. Yang J, Yan Y, Dai Q, Yin J, Zhao L, Li Y, et al. Tilorone confers robust in vitro and in vivo antiviral effects against severe fever with thrombocytopenia syndrome virus. *Virol Sin.* (2022) 37:145–8. doi: 10.1016/j.virs.2022.01.014

42. Małkowska P, Niedzwiedzka-Rystwej P. Factors affecting RIG-I-like receptors activation—new research direction for viral hemorrhagic fevers. *Front Immunol.* (2022) 13:1010635. doi: 10.3389/fimmu.2022.1010635

43. Ganesh GV, Mohanram RK. Metabolic reprogramming and immune regulation in viral diseases. *Rev Med Virol.* (2021) 32:e2268. doi: 10.1002/rmv.2268

44. Seong GM, Hyun CL, Chang SW, Yoo JR. Proven nasal and pulmonary aspergillosis in patient with severe fever with thrombocytopenia syndrome. *Tuberc Respir Dis.* (2022) 85:276–8. doi: 10.4046/trd.2022.0030

45. Teng A-Y, Che T-L, Zhang A-R, Zhang Y-Y, Xu Q, Wang T, et al. Mapping the viruses belonging to the order Bunyavirales in China. *Infect Dis Poverty.* (2022) 11:43–61. doi: 10.1186/s40249-022-00993-x

46. Wang T, Xu L, Zhu B, Wang J, Zheng X. Immune escape mechanisms of severe fever with thrombocytopenia syndrome virus. *Front Immunol.* (2022) 13:937684. doi: 10.3389/fimmu.2022.937684

47. Lan X, Zhang Y, Jia X, Dong S, Liu Y, Zhang M, et al. Screening and identification of Lassa virus endonuclease-targeting inhibitors from a fragment-based drug discovery library. *Antiviral Res.* (2021) 197:105230. doi: 10.1016/j.antiviral.2021.105230

frontiers | Frontiers in Oncology

TYPE Opinion
PUBLISHED 22 February 2023
DOI 10.3389/fonc.2023.1074091

OPEN ACCESS

EDITED BY
Yingying Xu,
Beihang University, China

REVIEWED BY
Christian Chabannon,
Aix-Marseille Université, France

*CORRESPONDENCE
Richard J. Epstein
r.epstein@unsw.edu.au

SPECIALTY SECTION
This article was submitted to Cancer Epidemiology and Prevention, a section of the journal Frontiers in Oncology

RECEIVED 01 December 2022
ACCEPTED 17 January 2023
PUBLISHED 22 February 2023

CITATION
Epstein RJ, Gu Y and Lin FPY (2023) Can cancer go green? It's up to us. *Front. Oncol.* 13:1074091. doi: 10.3389/fonc.2023.1074091

Can cancer go green? It's up to us

Richard J. Epstein[1,2,3]*, Yanfei Gu[1] and Frank P. Y. Lin[2,3,4]

[1]New Hope Cancer Center, Beijing United Hospital, Beijing, China, [2]Department of Medicine, University of New South Wales, Sydney, NSW, Australia, [3]Cancer Programme, Garvan Institute of Medical Research, Sydney, NSW, Australia, [4]National Health & Medical Research Council (NHMRC) Clinical Trials Centre, Sydney University, Sydney, NSW, Australia

KEYWORDS

oncology, health economics, sustainability, demographics, population aging

Introduction

The problem of cancer has long been supported by public taxation, private philanthropy and business investment (1), with this support having been both a cause and effect of clinical and scientific breakthroughs – including but not limited to adjuvant therapies, targeted therapies, immunotherapies, digitised imaging technologies, genetic sequencing, and cancer-preventive vaccines. Indeed, the term "oncology" only entered professional usage after proclamation of a War on Cancer by President Nixon in the lead-up to his 1972 re-election (2); this anti-cancer campaign was revived in 2016 by the Cancer Moonshot project, "re-ignition" of which was declared in 2022 by President Biden, with the goal of reducing age-specific cancer deaths by 50% over 25 years (3).

To understand the success [notwithstanding certain caveats (4)] of what has so far been a half-century campaign, it should first be asked why cancer has attracted more funding per unit of disease-specific mortality than have most other health issues; for example, there has been no similar support for a War on Heart Disease, even though cardiovascular problems have long caused higher death rates and health costs than cancer (5). This depth of support for cancer has been attributed to perceptions that a cancer diagnosis presents a unique existential threat [i.e., an "unspeakable" illness (6)] that not only poses lethal risks but also creates spiritually arduous – whether "moralistic" or "militaristic" – uncertainties as to the timing of cancer recurrence, nature of future symptoms, disease response or resistance, and speed (rapidity or slowness) of death (7).

Nonetheless, since resources are limited in any system, even the most serious personal health concerns (8, 9) must ultimately compete for support with societal-level threats (10, 11). During the first two decades (1971-90) of the War on Cancer, such threats included overpopulation, risks of nuclear war, and the HIV pandemic. Concerns over these issues abated over time; in hindsight, this fading of competing threats – akin to a peace dividend (12) – enabled sharper focus on the individual risk of a cancer diagnosis. Looking ahead, although cancer will always loom large as a major worry for personal health (13), the association of this disease with aging [i.e., with a low detriment to species fertility (14)] ensures that its impact on humanity will remain modest (15).

End of the golden age

Competing threats to the traditional support base for cancer lie ahead. Amongst these are the approaching impacts of global population aging, in tandem with falling birth rates (16), on oncology practice. Population aging will increase the aggregate burden of cancer diagnoses (17), even as age-specific cancer prevalence declines due to preventive advances; hence, as populations age over the coming decades, cancer will become commoner, but mainly among adults older than 65 years who will tend to have more age-related frailties than those diagnosed (younger) in the past (18). On the positive side, this greater longevity partly reflects improved disease prevention and wellness (19) ("healthy aging", "delayed aging"), just as falling fertility may arise to some extent from more effective contraception (20).

Whatever the reasons for population aging, a key consequence for today's oncologists is that their future (older) patients – who despite healthy aging are likely to have on average more restricted activities of daily living, more tenuous quality of life, and more competing causes of death (18) – may come to value autonomy and life quality relatively more highly than did their survival-focused predecessors (21, 22). The advent of patient-reported outcome measures represents a major step forward in this process of change, signaling as it does a 'personalisation' not of treatment targeting but of quality-of-life feedback and optimisation. Although such changes will not transform practice in the present decade, by 2040 evidence of this transition is predicted to become clear (23).

A different threat to human livelihoods is environmental degradation (24). There may seem little that links oncology practice and the causes or effects of environmental decline; yet when all correlates of the latter problem are considered – global warming, weather disasters, species displacements driving risks of disease spread to humans, travel restrictions, decarbonisation costs, energy shortages, rising prices, geopolitical instability – a total disconnect seems unlikely (25). At the least, traditional access to funding and philanthropy is likely to be constrained by competing cost-intensive problems like climate change (26), with the result – amplified by population aging – that resources for cancer become squeezed. Hence, although paradigm-shifting therapeutic advances must continue to add benefit for cancer patients in the future, other emerging challenges are likely to diminish investment in cancer therapeutics over time, especially when expressed as a fraction of world expenditures against all threats.

Public acceptance of this new reality could bring with it a slow change from an individualistic view of health as priceless (27) to a more socially cognizant "doughnut economics" (28), with proposals of this kind having already been made (29). Such a trend is also consistent with the Moonshot initiative which, in contrast to the War on Cancer, is not driven by major new funds; instead, the Moonshot proposes a switch of emphasis limited to an additional 5% of existing funding, with prevention as the priority. A transition of this kind aligns with recommendations to invest in a Culture of Health (30), based on a mindset valuing shared community needs (31, 32) at least as highly as the 'magic bullet' hope and hype which has for so long energised cancer research (33).

Steps to a greening of oncology

Different cultures of oncology already exist in different parts of the world, implying that changes are possible in any knowledge system (34). For cancer care to evolve from its golden age values to a more equable and lower-profile green age culture (Table 1), resistance may be eased by educational campaigns which can convince both public and professionals that these selection pressures reflect essential adaptive challenges for modern healthcare (35). A stepwise approach to this cultural transition is suggested below.

1. Win the battle of ideas over good-enough cancer treatment

Change in oncology is slowed by the idea that drug treatment choice – whether right [i.e., best, usually the latest [36)] or wrong (anything other than the best) – is a main determinant of disease outcome. As shown by the small absolute size of most clinical trial benefits, however, the reality is that cancer biology still tends to be the main factor affecting outcomes, with choices between approved

TABLE 1 Comparison of variables proposed to differ between the past (golden age) and potential future (green age) of oncology.

Variable	Golden age	Green age
Public/professional outlook	Only the best cancer treatment is good enough	Good enough cancer treatment may be the best
Clinical research study design	Largely specified by drug companies	Involves payers and patient representatives
Preferred site of standard care	Comprehensive Cancer Center	Community Cancer Center
Cancer patient management	Led throughout by Medical Oncologist	Shared from the start amongst MDT* members
Clinical decision-making	Based mainly on drug trial survival data	Based on nuanced patient-centric algorithms**
Attention to quality of life	Late, reactive to symptoms, 'damage control'	Early, anticipatory, preventive, maintenance
Patient mental health care	As needed, often Psychiatry-based	Pro-active, initially Psychology-based

*MDT, multidisciplinary team meeting, including at least one Medical Oncologist, plus additional Oncologists (Medical, Radiation, and/or Surgical), and at least one Supportive & Palliative Care specialist, and/or Pain Care specialist, Gerontologist, plus at least one Psychiatrist/Psychologist, Oncology Nurse Practitioner (including Stoma Care, Breast Care, etc.), Dietitian, Social Worker, Physiotherapist/Rehabilitation/Exercise Physiologist. **Including not only drug trial survival data (adjusted for level of evidence, statistical power, interpretability of study design, prospective vs. retrospective, primary endpoints, overall vs. disease-free survival, presence or absence of patient crossover, etc.), short- and long-term drug toxicity and tolerability, and quality of life, as per ASCO Value Framework and ESMO-MCBS (see text for references); plus financial costs, both absolute and out-of-pocket; patient convenience; patient autonomy; and other patient preferences; as predicted with or without a given treatment, and compared with other options, including no treatment.

treatments [or differences in national expenditures thereon (37)] usually making only modest differences (38). The modern incrementalist culture may exploit patients' fear of death (39) by perpetuating the norm that only the best is good enough (40), leading to the paradox that a disproportionate amount of oncology costs are generated in the last months of life (41). Such thinking may reflect a survival-of-the-fittest instinct, whereby any threat triggers a safety-first (kill it before it kills you) reaction; yet in the real world less extreme responses may not only suffice, but also prove less morbid (42).

2. Re-prioritise from research-centered to patient-centered practice

Focusing solely on the disease-modifying potential of therapies – with the potentially gratifying but often undetectable benefits of immune checkpoint therapy being a case in point (43) – distracts from the importance of factoring in other decision-making criteria such as treatment tolerability, safety, convenience, cost, options of delaying treatment (e.g., by later sequential or crossover drug use), and so on (44). One way to break this cycle, and thus to assess the validity or otherwise of good-enough treatment, is to quantify all factors pertinent to a cancer patient's predicted length and quality of life with or without the prescription of a treatment (45). If this exercise becomes possible to score – a technical quantum leap not yet reduced to practice – a good-enough cancer therapy could, paradoxically, deliver superior overall (i.e., holistic) outcomes compared to the best cancer treatment as determined by survival data. One step towards this has been the use of clinical trial or gene expression data to predict when addition of adjuvant cytotoxic therapy to hormone therapy delivers such a low absolute breast cancer survival increment as to be not recommendable as a standard of care (46). Validated algorithms which can quantify a given patients' preferences and priorities thus seem a prerequisite for progress.

3. Abandon zero-sum mindsets

Career success in newer subspecialties, such as oncology, has long been favoured by a tight focus on established (homophilic) goals and colleagues (47) – the so-called silo effect – rather than by more collaborative approaches (48). Yet if it is assumed that the problems of oncology do not overlap with those of other important contextual issues, such as the aging population or environmental deterioration, then a zero-sum interaction is assured; that is, any competition for resources between these fields will proceed on an 'I-win-you-lose' basis. In contrast, if common goals can be discerned, win-win scenarios for mutual benefit, and hence for co-resourcing, could be pursued (49, 50). Examples of integrated initiatives for cancer patients and the public include:

- Smoking elimination, clean air, ambient toxin reduction (e.g., radon, asbestos)
- Addiction prevention, mental resilience promotion
- Vaccination drives (e.g., HPV, HBV)
- Exercise programs (51)
- Ideal body weight education
- Food/beverage labeling improvements
- More multidisciplinary hospital-based care
- More community-based care for standard clinical problems
- User-friendly software development to aid more nuanced decision-making

4. Question self-reinforcing feedback loops

A risk of any golden age is that it selects for its own survival. In oncology, therapeutic progress has always been similarly sought by physicians, patients, pharmas, philanthropists, and the press; the only stakeholders who are motivated to query such progress are the third-party payers (52). This near-unanimity over the desirability of constant progress has made objective debate difficult (53), with the careers of critics prone to damage (54). The best solution to this problem will be to develop and validate quantitative metrics that can value holistic (i.e., qualitative) patient-centered variables, thus extending and refining the meaning of cost-efficacy (55).

5. Broaden decision-making and management

This quest to validate predictors of therapeutic value has made progress with upgradings of both the American Society of Clinical Oncology (ASCO) Value Framework Net Benefit Score (56) and the European Society of Medical Oncology (ESMO) Magnitude of Clinical Benefit Scale (MCBS) (57). The pros and cons of these sophisticated tools have been compared (58), and reveal promising complementarity in the questions addressed (59). Implementation of these algorithms remains labour-intensive, however, hampering patient inputs and clinical adoption (60). Broadening management to include routine early involvement of non-oncology multidisciplinary experts should help to dilute what is now, for many standard-setting cancer patients, an overly specialised approach to care (61). Better patient experiences could also result from moving standard care away from specialised cancer centers into general hospitals or local communities, assisted where needed by telehealth communications, while also evolving towards value-based reimbursement systems (62). More promotion of oncologists for excellence in communication or education (63) could be another step towards a flatter service structure. Finally, more attention to the mental resilience of patients, will also add value to patients' well-being, in part by reducing reliance on test results as critical arbiters of survivorship (64).

Conclusion

The choice of a War on Cancer as an encore to the Apollo moon landings signalled the zenith for cancer as an existential human

threat, kicking off a golden age for the science and practice of oncology. Times have since changed, however, as population aging and planetary threats have altered the trajectory of public concerns. The field of oncology is likely to feel these pressures to adapt within the next two decades, and the healthcare changes that result over the next fifty years could prove to be just as important and beneficial as the paradigm shifts that preceded these.

A global approach, based on public and professional education (65), will be needed to bring about economics-based system changes that can adapt to the disruptive evolutionary era ahead (66). Crosstalk with all stakeholders – including, though not limited to, patients, physicians, advocacy groups, governments, insurers, and pharmas – will necessarily precede such changes. The 20th century paradigm of ever more resources (ultimately derived from the environment) being harnessed to deliver ever more personalised oncologic increments (ultimately benefiting individuals more than populations) may prove to be less sustainably applicable to the 21st century world, where an imbalance between human demand and ecosystem fragility has become evident. Constructive change will require that concepts such as 'greater good', 'big picture' and 'longer term' come to be pursued more systematically than the self-interested priorities of the past.

Author contributions

RE wrote the first draft with inputs from FL and YG, then the second draft was further appraised and critiqued by FL and YG. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Sam D, Cheung WY. A population-level comparison of cancer-related and non-cancer-related health care costs using publicly available provincial administrative data. *Curr Oncol* (2019) 26:94–7. doi: 10.3747/co.26.4399

2. Kennedy BJ. Medical oncology: The new subspecialty. *Med Pediatr Oncol* (1975) 1:23–7. doi: 10.1002/mpo.2950010105

3. Ledford H. Cancer 'moonshot' has lofty new goal: Halve deaths in 25 years. *Nature* (2022) 602:561. doi: 10.1038/d41586-022-00376-0

4. Sonnenschein C, Soto AM. Over a century of cancer research: Inconvenient truths and promising leads. *PloS Biol* (2020) 18:e3000670. doi: 10.1371/journal.pbio.3000670

5. Muka T, Imo D, Jaspers L, Colpani V, Chaker L, van der Lee SJ, et al. The global impact of non-communicable diseases on healthcare spending and national income: A systematic review. *Eur J Epidemiol* (2015) 30:251–77. doi: 10.1007/s10654-014-9984-2

6. Pierotti MA. The molecular understanding of cancer: from the unspeakable illness to a curable disease. *Ecancermedicalscience* (2017) 11:747. doi: 10.3332/ecancer.2017.747

7. Curran L, Sharpe L, MacCann C, Butow P. Testing a model of fear of cancer recurrence or progression: The central role of intrusions, death anxiety and threat appraisal. *J Behav Med* (2020) 43:225–36. doi: 10.1007/s10865-019-00129-x

8. Hood B. *The self illusion: How the social brain creates identity*. New York: Oxford University Press (2012).

9. Allen SE. Cancer survivors: The success story that's straining health care. *IEEE Pulse* (2017) 8:14–7. doi: 10.1109/MPUL.2016.2629999

10. Badulescu D, Simut R, Badulescu A, Badulescu AV. The relative effects of economic growth, environmental pollution and non-communicable diseases on health expenditures in European union countries. *Int J Environ Res Public Health* (2019) 16:5115. doi: 10.3390/ijerph16245115

11. Bae EY, Lim MK, Lee B, Bae G. Who should be given priority for public funding? *Health Policy* (2020) 124:1108–14. doi: 10.1016/j.healthpol.2020.06.010

12. Al Mandhari A, Ghaffar A, Etienne CF. Harnessing the peace dividends of health. *BMJ Glob Health* (2021) 6:e006287. doi: 10.1136/bmjgh-2021-006287

13. Vehling S, Kissane DW. Existential distress in cancer: Alleviating suffering from fundamental loss and change. *Psychooncology* (2018) 27:2525–30. doi: 10.1002/pon.4872

14. Greaves M. *Cancer: The evolutionary legacy*. Oxford University Press. (2000)

15. Jasienska G, Bribiescas RG, Furberg AS, Helle S, Nunez-de la Mora A. Human reproduction and health: an evolutionary perspective. *Lancet* (2017) 390:510–20. doi: 10.1016/S0140-6736(17)30573-1

16. Pantazis A, Clark SJ. A parsimonious characterization of change in global age-specific and total fertility rates. *PLoS One* (2018) 13:e0190574. doi: 10.1371/journal.pone.0190574

17. Nolen SC, Evans MA, Fischer A, Corrada MM, Kawas CH, Bota DA. Cancer-incidence, prevalence and mortality in the oldest-old. a comprehensive review. *Mech Ageing Dev* (2017) 164:113–26. doi: 10.1016/j.mad.2017.05.002

18. Gu YF, Lin FP, Epstein RJ. How aging of the global population is changing oncology. *Ecancermedicalscience* (2021) 15:ed119. doi: 10.3332/ecancer.2021.ed119

19. Kemoun P, Ader I, Planat-Benard V, Dray C, Fazilleau N, Monsarrat P, et al. A gerophysiology perspective on healthy ageing. *Ageing Res Rev* (2022) 73:101537. doi: 10.1016/j.arr.2021.101537

20. Gotmark F, Andersson M. Human fertility in relation to education, economy, religion, contraception, and family planning programs. *BMC Public Health* Oxford UK (2020) 20:265. doi: 10.1186/s12889-020-8331-7

21. Gawande A. *Being mortal: medicine and what matters in the end*. New York: Metropolitan Books (2014).

22. Philipp R, Mehnert A, Lo C, Muller V, Reck M, Vehling S. Characterizing death acceptance among patients with cancer. *Psychooncology* (2019) 28:854–62. doi: 10.1002/pon.5030

23. Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet* (2018) 392:2052–90. doi: 10.1016/S0140-6736(18)31694-5

24. Sage RF. Global change biology: A primer. *Glob Chang Biol* (2020) 26:3–30. doi: 10.1111/gcb.14893

25. Frumkin H, Haines A. Global environmental change and noncommunicable disease risks. *Annu Rev Public Health* (2019) 40:261–82. doi: 10.1146/annurev-publhealth-040218-043706

26. Iyer HS, DeVille NV, Stoddard O, Cole J, Myers SS, Li H, et al. Sustaining planetary health through systems thinking: Public health's critical role. *SSM Popul Health* (2021) 15:100844. doi: 10.1016/j.ssmph.2021.100844

27. Poundstone W. *Priceless: The myth of fair value*. New York: Hill & Wang (2010).

28. Raworth K. *Doughnut economics: 7 ways to think like a 21st-century economist*. Vermont: Chelsea Green (2017).

29. Farrelly C. Responsible biology, aging populations and the 50th anniversary of the "War on cancer". *Biogerontology* (2021) 22:429–40. doi: 10.1007/s10522-021-09925-y

30. Tziraki-Segal C, De Luca V, Santana S, Romano R, Tramontano G, Scattola P, et al. Creating a culture of health in planning and implementing innovative strategies addressing non-communicable chronic diseases. *Front Sociol* (2019) 4:9. doi: 10.3389/fsoc.2019.00009

31. Prasad V. Our best weapons against cancer are not magic bullets. *Nature* (2020) 577:451. doi: 10.1038/d41586-020-00116-2

32. Marrauld L, Rambaud T, Egneli M, Geist JN. *THESHIFTPROJECT, « DÉCARBONER LA SANTÉ POUR SOIGNER DURABLEMENT »: LE RAPPORT DU SHIFT PROJECT*. Paris, France: The Shift Project (2021).

33. Sanchez NS, Mills GB, Mills Shaw KR. Precision oncology: Neither a silver bullet nor a dream. *Pharmacogenomics* (2017) 18:1525–39. doi: 10.2217/pgs-2017-0094

34. Haynes A, Rychetnik L, Finegood D, Irving M, Freebairn L, Hawe P. Applying systems thinking to knowledge mobilisation in public health. *Health Res Policy Syst* (2020) 18:134. doi: 10.1186/s12961-020-00600-1

35. Hensher M, Tisdell J, Canny B, Zimitat C. Health care and the future of economic growth: exploring alternative perspectives. *Health Econ Policy Law* (2020) 15:419–39. doi: 10.1017/S1744133119000276

36. Danzon PM. Drug pricing and value in oncology. *Recent Results Cancer Res* (2019) 213:153–67. doi: 10.1007/978-3-030-01207-6_10

37. Hermanowski T, Bystrov V, Staszewska-Bystrova A, Szafraniec-Burylo SI, Rabczenko D, Kolasa K, et al. Analysis of trends in life expectancies and per capita gross domestic product as well as pharmaceutical and non-pharmaceutical healthcare expenditures. *Acta Pol Pharm* (2015) 72:1045–50.

38. Tannock I, Presley CJ, Saltz LB. Value-added decisions in oncology. *Am Soc Clin Oncol Educ Book* (2019) 39:122–31. doi: 10.1200/EDBK_238831

39. Menzies RF, Menzies RG. Mortals: How the fear of death shaped human society. Sydney: Allen & Unwin (2021).

40. Engelberg AB, Avorn J, Kesselheim AS. A new way to contain unaffordable medication costs - exercising the government's existing rights. *N Engl J Med* (2022) 386:1104–6. doi: 10.1056/NEJMp2117102

41. Dvortsin E, Gout-Zwart J, Eijssen EL, van Brussel J, Postma MJ. Comparative cost-effectiveness of drugs in early versus late stages of cancer; review of the literature and a case study in breast cancer. *PloS One* (2016) 11:e0146551. doi: 10.1371/journal.pone.0146551

42. Milo DS. *Good enough: The tolerance for mediocrity in nature and society*. Cambridge, MA USA: Harvard University Press (2019).

43. Pichler M, Steyrer J. Cost-effectiveness analysis of the use of immunotherapy in metastatic solid tumours in Austria by applying the ESMO-magnitude of clinical benefit scale (ESMO-MCBS) version 1.1. *ESMO Open* (2021) 6:100198. doi: 10.1016/j.esmoop.2021.100198

44. Schnog JB, Samson MJ, Gans ROB, Duits AJ. An urgent call to raise the bar in oncology. *Br J Cancer* (2021) 125:1477–85. doi: 10.1038/s41416-021-01495-7

45. Walsh S, de Jong EEC, van Timmeren JE, Ibrahim A, Compter I, Peerlings J, et al. Decision support systems in oncology. *JCO Clin Cancer Inform* (2019) 3:1–9. doi: 10.1200/CCI.18.00001

46. Caparica R, Brandao M, Piccart M. Systemic treatment of patients with early breast cancer: Recent updates and state of the art. *Breast* (2019) 48 Suppl 1:S7–S20. doi: 10.1016/S0960-9776(19)31115-4

47. Woelmer WM, Bradley LM, Haber LT, Klinges DH, Lewis ASL, Mohr EJ, et al. Ten simple rules for training yourself in an emerging field. *PloS Comput Biol* (2021) 17:e1009440. doi: 10.1371/journal.pcbi.1009440

48. Akerlof KL. Beyond the sheltering academic silo: Norms for scientists' participation in policy. *Prog Mol Biol Transl Sci* (2022) 188:29–44. doi: 10.1016/bs.pmbts.2021.11.007

49. Malhi Y, Franklin J, Seddon N, Solan M, Turner MG, Field CB, et al. Climate change and ecosystems: Threats, opportunities and solutions. *Philos Trans R Soc Lond B Biol Sci* (2020) 375:20190104. doi: 10.1098/rstb.2019.0104

50. Bretti S, Porcile G, Romizi R, Palazzo S, Oliani C, Crispino S, et al. "Green oncology": The Italian medical oncologists' challenge to reduce the ecological impact of their clinical activity. *Tumori* (2014) 100:e94–7. doi: 10.1177/1578.17246

51. Kennedy MA, Bayes S, Newton RU, Zissiadis Y, Spry NA, Taaffe DR, et al. Implementation barriers to integrating exercise as medicine in oncology: An ecological scoping review. *J Cancer Surviv* (2022) 16:865–81. doi: 10.1007/s11764-021-01080-0

52. Nabhan C, Phillips EG, Feinberg BA. Value in oncology: It is in the eyes of the beholder. *J Natl Compr Canc Netw* (2019) 17:2–5. doi: 10.6004/jnccn.2018.7092

53. Hordern J, Maughan T, Feiler T, Morrell L, Horne R, Sullivan R. The 'molecularly unstratified' patient: A focus for moral, psycho-social and societal research. *BioMed Hub 2* (2017) 2(suppl. 1):146–53. doi: 10.1159/000480422

54. Harris R. Tweeting oncologist draws ire and admiration for calling out hype. *Shots: Health News NPR* (2018).

55. Grossmann N, Wolf S, Rothschedl E, Wild C. Twelve years of European cancer drug approval-a systematic investigation of the 'magnitude of clinical benefit'. *ESMO Open* (2021) 6:100166. doi: 10.1016/j.esmoop.2021.100166

56. American Society of Clinical Oncology. *ASCO value in cancer care* (2022). Available at: https://www.asco.org/news-initiatives/current-initiatives/cancer-care-initiatives/value-cancer-care.

57. Kiesewetter B, Cherny NI, Boissel N, Cerisoli F, Dafni U, de Vries EGE, et al. EHA evaluation of the ESMO-magnitude of clinical benefit scale version 1.1 (ESMO-MCBS v1.1) for haematological malignancies. *ESMO Open* (2020) 5:e000611. doi: 10.1136/esmoopen-2019-000611

58. Saluja R, Everest L, Cheng S, Cheung M, Chan KKW. Assessment of whether the American society of clinical oncology's value framework and the European society for medical oncology's magnitude of clinical benefit scale measure absolute or relative clinical survival benefit: An analysis of randomized clinical trials. *JAMA Oncol* (2019) 5:1188–94. doi: 10.1001/jamaoncol.2019.0818

59. Cherny NI, de Vries EGE, Dafni U, Garrett-Mayer E, McKernin SE, Piccart M, et al. Comparative assessment of clinical benefit using the ESMO-magnitude of clinical benefit scale version 1.1 and the ASCO value framework net health benefit score. *J Clin Oncol* (2019) 37:336–49. doi: 10.1200/JCO.18.00729

60. Gyawali B, de Vries EGE, Dafni U, Amaral T, Barriuso J, Bogaerts J, et al. Biases in study design, implementation, and data analysis that distort the appraisal of clinical benefit and ESMO-magnitude of clinical benefit scale (ESMO-MCBS) scoring. *ESMO Open* (2021) 6:100117. doi: 10.1016/j.esmoop.2021.100117

61. Tahara DC, Green RP. Strategic re-design of team-based patient-focused health care services. *Adv Health Care Manag* (2014) 16:3–22. doi: 10.1108/S1474-823120140000016000

62. Cox JV, Ward JC, Hornberger JC, Temel JS, McAneny BL. Community oncology in an era of payment reform. *Am Soc Clin Oncol Educ Book* (2014), e447–52. doi: 10.14694/EdBook_AM.2014.34.e447

63. Villaire M, Mayer G. Health literacy: The low-hanging fruit in health care reform. *J Health Care Finance* (2009) 36:55–9.

64. Mayer DK, Nasso SF, Earp JA. Defining cancer survivors, their needs, and perspectives on survivorship health care in the USA. *Lancet Oncol* (2017) 18:e11–8. doi: 10.1016/S1470-2045(16)30573-3

65. Wardill HR, Cheung YT, Boltong A, Charalambous A, Koczwara B, Lustberg M, et al. 'Share your views'-international consultation informs a patient engagement strategy for the multinational association of supportive care in cancer. *Support Care Cancer* (2022) 30:9953–61. doi: 10.1007/s00520-022-07366-y

66. Gibbs JF, Newman A, Stefanacci RG. Value-based focused global population health management. *J Gastrointest Oncol* (2021) 12:S275–89. doi: 10.21037/jgo-2019-gi-10

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 24 February 2023
DOI 10.3389/fpubh.2023.1075277

OPEN ACCESS

EDITED BY
Dianqin Sun,
Erasmus Medical Center, Netherlands

REVIEWED BY
Daniel Michaeli,
Heidelberg University, Germany
Amir-Houshang Omidvari,
Erasmus Medical Center, Netherlands

*CORRESPONDENCE
Davood Khalili
✉ dkhalili@endocrine.ac.ir

SPECIALTY SECTION
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

RECEIVED 20 October 2022
ACCEPTED 23 January 2023
PUBLISHED 24 February 2023

CITATION
Jamshidi A, Daroudi R, Aas E and Khalili D (2023) A cost-effectiveness analysis of risk-based intervention for prevention of cardiovascular diseases in IraPEN program: A modeling study. *Front. Public Health* 11:1075277. doi: 10.3389/fpubh.2023.1075277

A cost-effectiveness analysis of risk-based intervention for prevention of cardiovascular diseases in IraPEN program: A modeling study

Amirparviz Jamshidi[1,2], Rajabali Daroudi[3], Eline Aas[2] and Davood Khalili[1,4]*

[1]Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [2]Department of Health Management and Health Economics, University of Oslo, Oslo, Norway, [3]Department of Health Management, Policy and Economics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran, [4]Department of Biostatistics and Epidemiology, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Background: IraPEN, a program developed in Iran based on the World Health Organization (WHO) package of essential noncommunicable (PEN) disease interventions for primary healthcare, was launched in 2015. Preventive interventions for cardiovascular diseases (CVDs) are based on the level of risk calculated using the WHO CVD risk chart.

Objective: The main objective of this study was to measure the potential cost-effectiveness (CE) of IraPEN preventive actions for CVD in comparison with the *status quo*.

Methods: A CE analysis from a healthcare perspective was conducted. Markov models were employed for individuals with and without diabetes separately. Based on the WHO CVD risk chart, four index cohorts were constructed as low (<10%), moderate (10%–19%), high (20%–29%), and very high risk (≥30%). Life years (LY) gained and quality-adjusted life years (QALY) were used as the outcome measures.

Results: The intervention yields an incremental cost-effectiveness ratio (ICER) of $804, $551, and −$44 per QALY for moderate, high, and very high CVD risk in groups without diabetes, respectively. These groups gained 0.69, 0.96, and 1.45 LY, respectively, from the intervention. The results demonstrated an ICER of $711, $630, −$42, and −$71 for low, moderate, high, and very high-risk groups with diabetes, respectively, while they gained 0.46, 1.2, 2.04, and 2.29 years from the intervention.

Conclusion: The IraPEN program was highly cost-effective for all CVD risk groups in the individuals without diabetes except the low-risk group. The intervention was cost-effective for all patients with diabetes regardless of their CVD risk. The results demonstrated that the IraPEN program can likely provide substantial health benefits to Iranian individuals and cost savings to the national healthcare provider.

KEYWORDS

cost effectiveness, cardiovascular risk factors, cardiovascular diseases, modeling, diabetes preventing program, CVD prevention, primary health care, WHO PEN

Background

Cardiovascular diseases (CVDs) are the leading cause of death worldwide and people die from CVDs more than any other causes. CVDs are considered a development issue as almost 75% of global CVD deaths occur in low- and middle-income countries. However, the majority of CVDs can be prevented by reducing the burden of risk factors (1).

Iran with an 80 million population in 2016, as an upper-middle-income developing country, has acquired many achievements in the public health sector during the past three decades. A well-balanced referral system within a broad PHC network, even in far stretches of villages, could provide access to healthcare for 95% of the community and control communicable diseases efficiently. This accomplishment resulted in a life expectancy of more than 75 years for men and more than 77 years for women. At present, the transition to chronic and noncommunicable diseases (NCDs) including CVDs, cancer, and mental disorders is the main problem in the health system. Based on the last report by the World Health Organization (WHO), NCDs are estimated to account for almost 80% of total deaths in Iran, while almost half of them (43%) are caused by CVDs. The comparison of Iranians' CVD mortality rate with other countries shows that not only it is substantially higher than high-income countries but also it is much higher than the countries in the region (2).

During the last two decades, many positive actions (e.g., public education, opportunistic finding of diabetic and hypertensive cases in network system, etc.) have been carried out to control NCDs in the country. Despite major achievements, NCDs and their subsequent burden have increased in the country (3).

In 2010, the WHO launched the package of essential noncommunicable (PEN) disease intervention for primary care in low-resource settings to deliver an adequate quality of care and, consequently, reduce the burden of these diseases in developing countries. WHO PEN has effective tools to facilitate early diagnosis and management of CVD, chronic respiratory diseases, diabetes, and cancer to prevent their upcoming morbidities and premature mortalities (e.g., stroke, myocardial infarction, renal failure, blindness, amputations, etc.) (4).

In 2015, IraPEN, an adaptation of WHO PEN, was launched as a part of the national Healthcare Reform Plan in Iran. Providing universal healthcare coverage and access to NCD prevention and treatment for all were the main goals of this reform. The first phase of the IraPEN program had been piloted in four cities, and the results were promising. In 2018, the program was expanded to all provinces of Iran. It was expected that at this phase, the program would cover up to four million people, and then based on the results, it would be expanded nationwide (4). Due to the IraPEN project size and its impact on the national healthcare budget, it is essential to make a detailed evaluation of these pilot enforcements to pave the way for IraPEN national implementation. Therefore, there is a need for an economic analysis that can estimate costs and effects as well as an incremental cost-effectiveness ratio (ICER) of the IraPEN program in comparison with the *status quo* (no prevention). Therefore, the main objective of this analysis is to measure the cost per life-year (LY) gained as well as the cost per quality-adjusted life years (QALY) gained by the IraPEN program.

Methods

This study evaluates the potential cost-effectiveness (CE) of the IraPEN program in comparison to the *status quo* through a health economic evaluation and the outcomes expressed in terms of QALY and LY gained for each CVD risk group. The target group of this analysis is all Iranian people aged older than 40 years and the evaluated intervention is the same as the recommended intervention of WHO PEN which is included screening, monitoring, and medications.

Model structure

Two separate Markov decision models were developed to compare the long-term costs and health benefits of the IraPEN program (primary CVD prevention) with the *status quo* (no prevention) in two distinct scenarios. In the base case scenario, individuals without diabetes were included, while patients with diabetes were included in the alternative scenario. Each Markov model has four health states with transitions between the states according to age, sex, and the CVD risk characteristics of participants (Figure 1). In contrast to the usual Markov models, which are structured based on cohorts with average profiles, we decided to categorize the individuals based on their CVD risks. As the intervention (treatment) varied according to CVD risk level, it is logical to model them separately. In this way, we can take into account their specific characteristics. Therefore, based on WHO/ISH CVD risk prediction charts for EMR B, four index cohorts were constructed (5). These hypothetical cohorts were used as a representative for individuals with low, moderate, high, and very high CVD risk profiles. The CVD risk state represents the starting point for all people who are 40 years old. It was assumed that people in this state may either remain in the same health state, move to the stroke state, or CHD (coronary heart disease) state, or die. As long as they are event-free, these individuals can stay in a healthy state, but after the first event, they move to the CHD or stroke state and stay there until their death.

In WHO/ISH CVD risk prediction charts, the CVD risk is calculated based on individuals' age and risk factors such as blood pressure, lipid profile, diabetes, and smoking status and categorized into the following five groups: below 10% (low-risk group), between 10 and 19% (moderate-risk group), between 20 and 29% (high-risk group), between 30 and 39%, and above 40% (very high-risk group). As the individuals in the two latter groups are treated the same, in the IraPEN program, whoever has a CVD risk above 30% is categorized as the very high-risk group.

Therefore, considering what was mentioned earlier, all the Iranians aged older than 40 years who did not have CHD or stroke events before were eligible for this program. According to the recent census (2016), 31.16% of Iranians were older than 40 years (6). By adding individuals aged older than 30 years with the aforementioned risk factors, we can conclude that this program is going to screen at least 25 million people yearly.

The healthcare perspective and a 40-year time horizon were adopted for this analysis. As the analysis is a comparison between

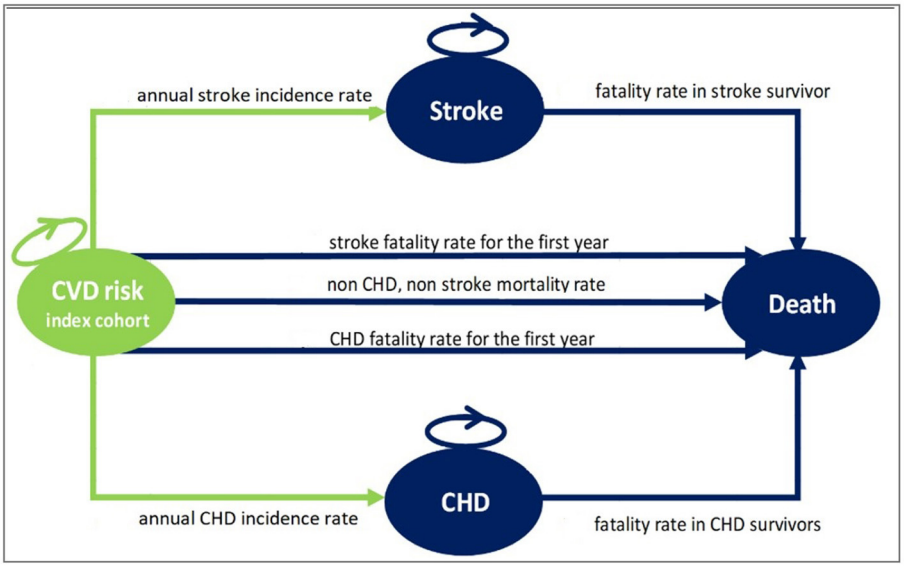


FIGURE 1
The structure of the Markov model used for the IraPEN analysis. CVD risk index cohort state, healthy individuals with different CVD risk; Stroke state, alive individuals after the first stroke event; CHD state, alive individuals after the first CHD event; Death state, dead individuals.

TABLE 1 Index cohorts representing different CVD risk levels*.

	Low risk	Moderate risk	High risk	Very high risk
Without diabetes				
Systolic blood pressure	120–139 mmHg	140–159 mmHg	160–179 mmHg	>180 mmHg
Total cholesterol	<195 mg/dl	>310 mg/dl	>310 mg/dl	>310 mg/dl
HDL	40 mg/dl	46 mg/dl	46 mg/dl	41 mg/dl
Smoking	No	Yes	Yes	Yes
Sex	Male	Female	Female	Male
With diabetes				
Systolic blood pressure	120–139 mmHg	140–159 mmHg	160–179 mmHg	>180 mmHg
Total cholesterol	<195 mg/dl	>310 mg/dl	<270 mg/dl	<270 mg/dl
HDL	46 mg/dl	46 mg/dl	40 mg/dl	41 mg/dl
Smoking	No	No	Yes	Yes
Sex	Female	Female	Male	Male

*Green color: represents the low-risk group (individuals with CVD risk <10%). Yellow color: represents the moderate-risk group (individuals with CVD risk between 10 and 19%). Orange color: represents the high-risk group (individuals with CVD risk between 20 and 29%). Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

IraPEN (intervention) and *status quo* (no intervention) which both have the same Markov structure and transition probabilities, it is not expected that half cycle correction (HCC) approach makes any difference in ICER results; therefore, HCC was not applied to this analysis (7).

The hypothetical cohorts were used as a representative for individuals with low, moderate, high, and very high CVD risk profiles (Table 1). Progressively, a proportion of the cohort can go to the CHD state, who are the survivors of the first CHD event, or to the stroke state who are the survivors of the first stroke event. Those CHD and stroke events that were fatal moved to the death state. In general, the people in these two states are at a higher risk of dying from CHD or stroke, but they may die from any other causes like the normal population. Table 2 summarizes the assumptions of this analysis.

Data input

This analysis tried to use the Iranian data wherever available. In case of a lack of local data, the inputs were derived from the global literature. Therefore, all transition probabilities of the models were extracted from available Iranian data, while medications' effects and

TABLE 2 Main elements of this economic evaluation (EE).

Comparators	IraPEN intervention vs. status quo (no prevention)
Perspective	Healthcare
Target group	All Iranians older than 40 years old
Type of EE	Cost-effectiveness analysis by adopting a Markov model
Considered costs	All direct medical costs
Discount rate	3.5% for costs and effects
Sensitivity analysis	Determinstic and Probabilistic sensitivity analysis (DSA and PSA)
CE threshold	GDP per capita of Iranian people—$4,091

states' utilities were driven from the literature of Western countries (refer to Supplementary material). No individual data have been used for this analysis.

Transition probabilities

The annual incidence rate for CHD (8) and stroke was calculated from the Framingham study equations (9). As four-index cohorts had been defined with specific characteristics, there was a need to calculate the risk based on those profiles. Based on the literature, one out of four CHD events are fatal in the first year (10), while 60% of them are pre-hospital deaths (11). Therefore, it was assumed that of those who have a CHD event in the model, 25% die in the first year. Approximately 60% of these deaths were costless as they are pre-hospital deaths. Regarding first-year stroke mortality, the range varies in different resources and is reported from 22 to 34% (12). For this analysis, the rate was applied from the largest available cohort (13). Almost 25% of stroke events are fatal in the first year, while half of them occur during the first 28 days. Therefore, it is assumed that although at the end of the cycle, they move to the death state, 40% of the cycle cost should be considered for them. The fatality rate for stroke and CHD survivors was derived from a study that had been done on the Iranian population (14). The background mortality rate from all causes other than stroke and CHD is calculated by excluding the total death attributed to these two diseases from the Iranian life table[1]. The total mortality rate of these two events had been calculated in Tehran Lipid and Glucose Study (TLGS). At first, the annual rates were derived from the life table and then the CHD- and stroke-attributed deaths were excluded.

Intervention effect

The IraPEN's preventive actions are expected to reduce cardiovascular events. The relative risks (RRs) of these preventive actions and the medications that are used in the program were obtained from meta-analyses or randomized clinical trials (RCTs). By multiplying or adding up the RRs of different medications, there is a risk of effect overestimation, and a correction was made by using the formula below wherever multiple interventions were involved:

$$1 - ((1 - RR_1 \times RR_2 \times RR_N) \times 0.8).$$

This equation has been developed based on a study that compared the effect of controlling the risk factors separately vs. controlling all of them simultaneously (15).

Based on the field interviews, it was clear which medications are used for each index cohort. Almost in all cases, angiotensin-converting enzyme (ACE) inhibitors are the first choice for hypertension treatment. Enalapril is the most prescribed one as monotherapy. Thiazides (diuretics) are the second choice followed by beta-blockers. In case the hypertension is not controlled by monotherapy instead of increasing the dose, the second drug is added. As recommended by guidelines, small doses of various classes of antihypertensive medications are more useful than a high dose of one (16). In general, the combination of ACE inhibitors and thiazide is the most common one. This pattern is aligned with Joint National Committee (JNC8) guidelines. Statins are prescribed for hyperlipidemia treatment. Among statins, Atorvastatin is the choice as it is one of the most potent ones. For diabetes, Metformin is started and increased to the maximum dose (2 g) and then the second medication that is Glibenclamide is added. Due to its potential harm and insufficient evidence of its efficacy, Aspirin was not recommended for primary prevention by PEN protocols. Therefore, Aspirin is not used in IraPEN as well. Here are the list of medications and their daily dosages which are used in IraPEN:

- Atorvastatin 20 mg tablet for statin therapy (statin).
- Enalapril Maleate 20 mg tablet is the first choice for hypertension treatment (Ace inhibitor).
- Hydrochlorothiazide 50 mg tablet (diuretics second choice).
- Metoprolol tartrate 50 mg tablet (beta-blocker third choice).
- Metformin HCL 500 mg tablet for diabetes (daily consumption from 500 to 2 g).
- Glibenclamide 5 mg tablet is the second choice for diabetes.

The unit price of each of these medications was derived from the Iranian Annual Pharma Statistics file. For the calculation of the intervention's effects, it is assumed that the adherence of individuals to the treatment is 100%. Table 3 lists the RRs of different interventions (medications) for CHD and stroke.

Costs

A healthcare perspective was adopted; therefore, we only included costs associated with healthcare such as direct medical costs (Table 4). The costs considered in the model are the cost of IraPEN screening, the cost of IraPEN monitoring, the cost of CHD survivors, and the cost of stroke survivors. It is assumed that the cost of individuals who are event-free in the *status quo* is zero as long as undiagnosed or untreated. These two facts were considered for the *status quo* costs. Furthermore, it was assumed that the cost of dying was equal to zero. According to

1 https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/gho-ghe-global-health-estimates-life-tables

PEN protocols, the needed resources for each index cohort were identified. Then, the items were quantified based on discussions with the physicians and supervisors of the visited centers. The cost of index cohorts consists of two different types. First, variable costs are different for each group based on the characteristics of each. Second, fixed cost is the same for all and consists of staff training, administration, IT, promotional stuff, and leaflets. The unit price of each item was derived from the last report of the Ministry of health (23). The report estimated all the costs related to IraPEN implementation except the medications. In addition, the reported costs were adjusted by the 2018 inflation rate and the cost of each cohort was calculated.

The cost of CHD state and stroke state was derived from an Iranian CE that had estimated the cost of these two states (24). These two costs contain all the related medical costs such as hospital admissions and procedures, monitoring, follow-ups, medications, and secondary prevention (Table 5). Based on experts' opinions, it is assumed that the cost of CHD after the first year would be a third and the cost of stroke state after the first year would be a quarter. In addition, it is assumed that the standard error of costs for the consecutive year is 10% of the mean.

TABLE 3 Intervention effects based on the subclass of medications that are used in IraPEN.

		RR	(95% CI)
Angiotensin-converting enzyme inhibitor (17)			
RR	CHD	0.81	(0.70–0.94)
	Stroke	0.65	(0.52–0.82)
Thiazide diuretics (18)			
RR	CHD	0.84	(0.75–0.95)
	Stroke	0.63	(0.57–0.71)
Beta-blockers (19)			
RR	CHD	0.90	(0.78–1.03)
	Stroke	0.83	(0.72–0.97)
Statin (20)			
RR	CHD	0.86	(0.82–0.90)
	Stroke	0.90	(0.85–0.95)
Metformin (21)			
RR	CHD	0.67	(0.51–0.89)
	Stroke	0.80	(0.50–1.27)
Sulfonylureas			
RR	CHD	0.85	(0.74–0.97)
	Stroke	0.91	(0.73–1.13)
Lifestyle counseling (22)			
RR	CHD	0.86	(0.81–0.91)
	Stroke	0.86	(0.81–0.91)

Utilities

As all people who enter the model are healthy individuals, the utility for the first health state is considered 1. For the death health state utility, it was adopted the standard approach by setting the utility to zero. CHD and stroke state utilities were derived

TABLE 5 Annual cost of CHD and stroke states.

	Mean	SE
CHD cost for the first year	$519	$51
CHD cost for consecutive years	$173	$17
Stroke cost for the first year	$5,691	$569
Stroke cost for consecutive years	$1,422	$142

TABLE 4 Annual cost of each index cohort with diabetes.

	Low-risk	Moderate-risk	High-risk	Very high-risk
Behvarz's visit (Screening)	108,173	324,519	432,692	432,692
Physician visit	–	216,346	288,462	432,692
Lab data (included in screening)		–	–	–
Nutrition consultation	38,400	38,400	38,400	38,400
Psychiatrist consultation	–	–	38,400	38,400
Anti-hypertensive medication	–	One agent	Two agents	Three agents
Statin Therapy	–	+	+	+
Fixed costs	19,231	19,231	19,231	19,231
Cost of each group in 2017 (IRR)	165,804	598,496	817,185	961,416
The inflation rate was applied to the costs (IRR)	216,540	781,636	1,067,243	1,255,609
Cost of each index cohort without medication	$5.16	$18.61	$25.41	$29.90
Cost of medication of each index cohort	$16	$38.85	$48.27	$55.39
Cost of each index cohort for the model	**$21.63**	**$57.46**	**$73.68**	**$85.29**

Exchange rate: 42,000 IRR = 1 US dollar. Bold indicates total cost of each index cohort for the model.
Green color: represents the low-risk group (individuals with CVD risk less than 10%); Yellow color: represents the moderate-risk group (individuals with CVD risk between 10% to 19%); Orange color: represents the high-risk group (individuals with CVD risk between 20% to 29%); Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

10.3389/fpubh.2023.1075277

TABLE 6 Utility weights for CHD and stroke states.

	Utility mean	SE
Utility of CHD survivors—first year	0.67	0.024
Utility of stroke survivors—first year	0.33	0.033
Utility of CHD survivors—second year onwards	0.82	0.012
Utility of stoke survivors—second year onwards	0.52	0.027

from the published literature. In the models after the first event, patients move to these states and stay there until they die. Although they remain in the same states (CHD or stroke), their utilities are different over time. From a medical view, acute post-event utilities are (much)lower than chronic post-event utilities. Therefore, it was essential to use the data from the study which assessed the acute and chronic utilities with the same participants at an appropriate time. For this purpose, the utilities were derived from the study which assessed the utilities within the first year and consecutive years (13). Table 6 shows the utilities used for the model. All the costs and effects were discounted at the rate of 3.5%.

Sensitivity analysis

To quantify the level of confidence in the models' results, a deterministic sensitivity analysis (DSA) and a probabilistic sensitivity analysis (PSA) were performed. In the DSA, the input parameters were varied to the maximum and minimum possible values. This range is usually defined by the confidence interval of parameters. Therefore, for the examined parameters, a range of 95% confidence interval was specified and then based on this range (maximum and minimum input values), one value in the model was varied manually each time. For the patients' adherence, the range of 50%−100% was considered. The results (new ICERs) were collected and expressed with tornado diagrams. The tornado diagrams depict the impact on the ICER whenever one single parameter changed.

For the PSA, as the main assumption, it was considered the deterministic input values in the parameter sheet as the mean values. As the standard errors of the cost items were not available, it was considered to mean value times by 0.1. Based on logical constraints, the probabilistic distribution for each of the different sources of uncertainty was defined. A gamma distribution for all cost items and a beta distribution for the utilities were defined. The PSA was conducted by drawing a random number for each of the input distributions and each time, the ICER was calculated by Excel. By running a macro, its action is repeated 1,000 times.

Results

For the interpretation of the results, 1 GDP per capita was assumed as the CE threshold, which is equal to $4,091 (25). In Table 7, the results of the "CVD risk model without diabetes" are reported. The IraPEN intervention for individuals having moderate CVD risk yields an ICER of $804 per QALY, while for high-risk groups, this intervention provides an ICER of $551 per QALY.

TABLE 7 Base-case results for CVD risk groups without diabetes.

	COST	QALY	LY	ICER
Status quo	$845	18.12	32.27	
IraPEN	$918	18.12	32.27	Undefined
Incremental	$73	0	0	Low-risk
Status quo	$979	17.89	32.37	
IraPEN	$1,375	18.38	33.06	$804 ♀
Incremental	$396	0.49	0.69	Intermediate-risk
Status quo	$1,204	17.68	32.13	
IraPEN	$1,594	18.38	33.09	$551 ♀
Incremental	$391	0.71	0.96	High-risk
Status quo	$2,348	15.83	29.3	
IraPEN	$2,296	17.02	30.75	−$44 ♂
Incremental	−$52	1.19	1.45	Very high-risk

CE threshold = $4,091, ♂ = male, ♀ = female.
Green color: represents the low-risk group (individuals with CVD risk less than 10%); Yellow color: represents the moderate-risk group (individuals with CVD risk between 10% to 19%); Orange color: represents the high-risk group (individuals with CVD risk between 20% to 29%); Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

TABLE 8 Base-case results for CVD risk groups with diabetes.

	COST	QALY	LY	ICER
Status quo	$647	18.59	33.36	Low risk
IraPEN	$866	18.9	33.82	$711 ♀
Incremental	$219	0.31	0.46	
Status quo	$1,027	17.75	32.17	Moderate
IraPEN	$1,560	18.59	33.37	$630 ♀
Incremental	$533	0.85	1.2	
Status quo	$2,429	15.91	29.46	High
IraPEN	$2,361	17.53	31.5	−$42 ♂
Incremental	−$68	1.61	2.04	
Status quo	$2,810	15.31	28.74	Very high
IraPEN	$2,678	17.18	31.03	−$71 ♂
Incremental	−$133	1.87	2.29	

CE threshold = $4,091, ♂ = male, ♀ = female.
Green color: represents the low-risk group (individuals with CVD risk less than 10%); Yellow color: represents the moderate-risk group (individuals with CVD risk between 10% to 19%); Orange color: represents the high-risk group (individuals with CVD risk between 20% to 29%); Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

The results showed that this intervention would be cost-saving and improve health if the IraPEN program targets only individuals with a CVD risk higher than 30%. The model yields the ICER of −$44 per QALY for this group. Moreover, the model results showed that individuals with higher CVD risks gained higher LY out of the intervention. The moderate, high, and very high CVD risk groups gained 0.69, 0.96, and 1.45 LY, respectively, from the

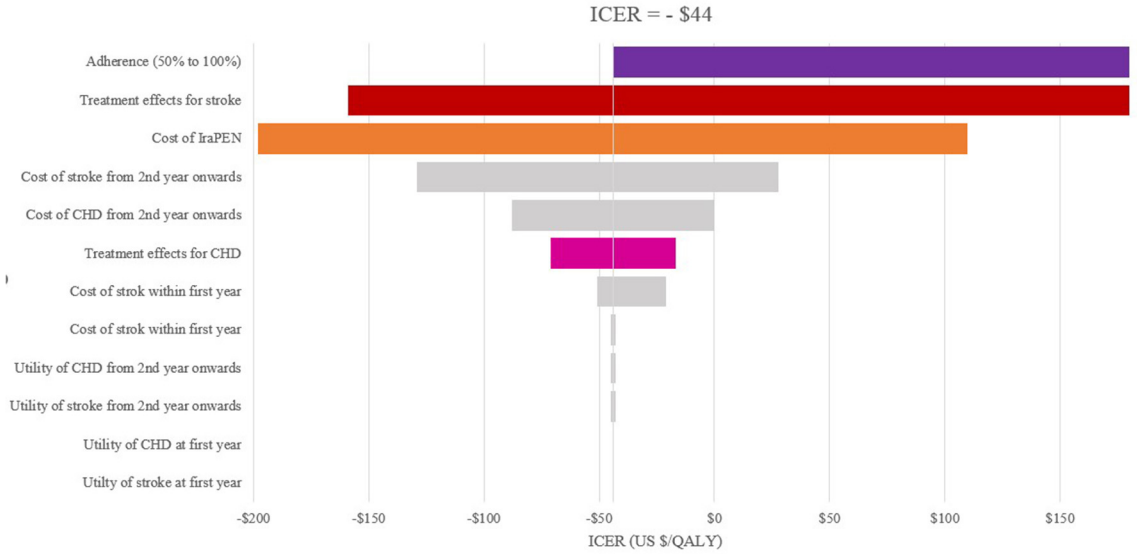


FIGURE 2
One-way sensitivity analysis of modeled cost–effectiveness of CVD risk-based prevention for very high-risk without diabetes group.

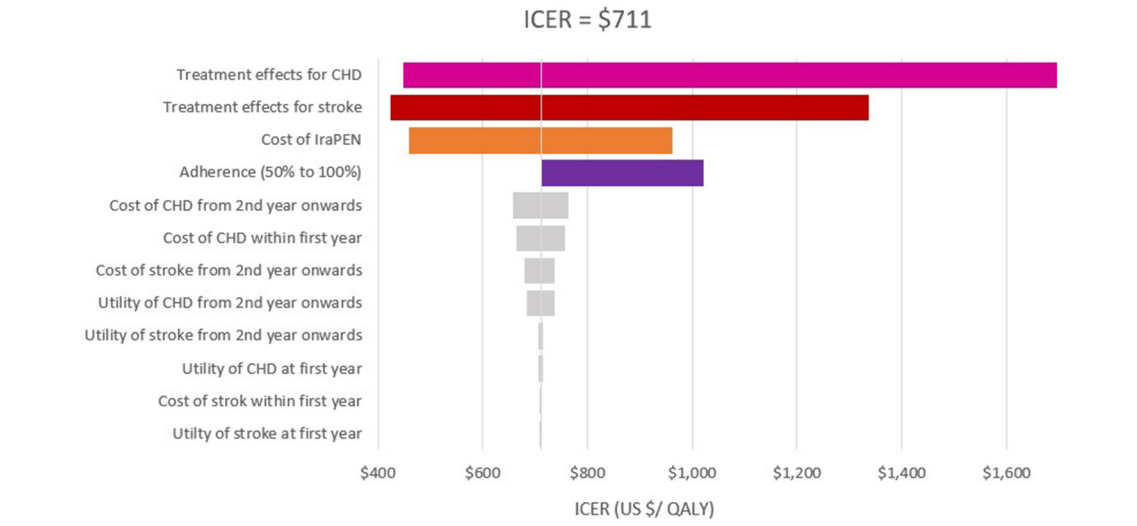


FIGURE 3
One-way sensitivity analysis of modeled cost–effectiveness of CVD risk-based prevention for low risk with diabetes group.

intervention. Table 8 shows the model results for the "CVD risk model with diabetes".

Here, the results demonstrate that the intervention would be cost-saving while improving health if target individuals with CVD risk comprise higher than 20%. The intervention yields an ICER of –$42 per QALY for the high-risk group and an ICER of –$71 per QALY for the very high-risk group. Similar to the previous model, individuals with higher risks gain more LY from the intervention.

The one-way sensitivity analysis reveals that the patients' adherence, the treatment effectiveness, and the total cost of the IraPEN program have the most impact on the ICER. The influence of patients' adherence is more noticeable in the higher CVD risk groups, while the results are more sensitive to treatment effectiveness in the lower risk groups (Figures 2, 3).

In the scenario analysis of 50% adherence, an ICER of $1,451, $1,141, and $329 was achieved for moderate, high, and very high CVD risk in groups without diabetes, respectively. In this scenario, the intervention yields an ICER of $1,029, $1,022, $236, and $199 for low, moderate, high, and very high CVD risk groups with diabetes, respectively. It was found that the ICERs are lower than the threshold at both the upper and lower limits of all examined parameters.

The result of 1,000 PSAs illustrates that the intervention for all groups, except the low CVD risk group without diabetes, is cost-effective while being cost-saving for at least half of high-risk

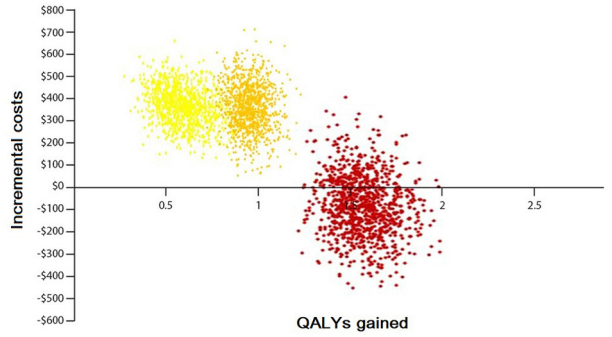


FIGURE 4
Probabilistic sensitivity analysis of the Markov model without diabetes. Yellow color: represents the moderate-risk group (individuals with CVD risk between 10 and 19%). Orange color: represents the high-risk group (individuals with CVD risk between 20 and 29%). Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

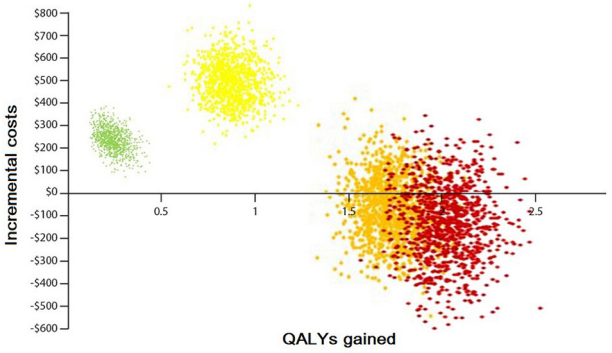


FIGURE 5
Probabilistic sensitivity analysis of the Markov model with diabetes. Green color: represents the low-risk group (individuals with CVD risk <10%). Yellow color: represents the moderate-risk group (individuals with CVD risk between 10 and 19%). Orange color: represents the high-risk group (individuals with CVD risk between 20 and 29%). Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

and very high-risk patients (Figures 4, 5). By adopting 1 GDP per capita of Iran as the willingness to pay per quality-adjusted life-year (WTP/QALY) gained, 100% of all the runs were cost-effective in these groups.

Both models captured that men would be benefited more than women in terms of LY gained. In addition, the intervention generates a lower ICER for men than women in individuals with identical characteristics. For example, in the low CVD risk group with diabetes, the interventions yield an ICER of $239 per QALY for men, while it is $711 per QALY for women. Regarding the LY in this group, in equal circumstances, men saved 0.58 of a year and women saved 0.46 of a year. In higher risk groups, this difference is more prominent. For example, for very high-risk groups, regardless of their diabetes status, the intervention for men is cost-saving, while for women it is not (Tables 9, 10).

TABLE 9 Heterogeneity base-case results for very high CVD risk group with diabetes.

	COST	QALY	LY		ICER
Status Que	$2,810	15.31	28.74	♂	
IraPEN	$2,678	17.18	31.03	♂	–$71
Incremental	-$133	1.87	2.29	♂	
Status Que	$2,133	16.17	30.18	♀	
IraPEN	$2,393	17.9	32.42	♀	$151
Incremental	260	1.73	2.24	♀	

CE threshold = $4,091, ♂ = male, ♀ = female.
Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

TABLE 10 Heterogeneity base-case results for very high CVD risk group without diabetes.

	COST	QALY	LY		ICER
Status Que	$2,348	15.83	29.3	♂	
IraPEN	$2,296	17.02	30.75	♂	–44
Incremental	-$52	1.19	1.45	♂	
Status Que	$1,522	17.19	31.46	♀	
IraPEN	$1,912	18.13	32.71	♀	418
Incremental	390	0.93	1.25	♀	

CE threshold = $4,091, ♂ = male, ♀ = female.
Red color: represents the very high-risk group (individuals with CVD risk more than 30%).

Discussion

This analysis aimed to measure the potential CE of IraPEN preventive actions for CVD in comparison with the *status quo*. Our results illustrated that this intervention is not cost-effective for the low CVD risk group, whereas the other groups under study proved to be highly cost-effective. The reason why this group was not cost-effective could be justified by the fact that the low CVD risk group has lower CVD risk factors, which is why such individuals are just screened without being intervened.

In this study, we adopted the CE threshold, i.e., 1–3 GDP per capita, as proposed by WHO (26). The study of the literature shows that by using this threshold, almost all interventions seem to be cost-effective (27). It means that by adopting this CE threshold, there is a risk that the budgets are spent on interventions that should not and vice versa. The threshold which is recommended by WHO has received some criticism as it is believed that it does not reflect the true "opportunity cost." This is more critical in low- to middle-income countries, because while they have a higher demand for health, in comparison with high-income countries, fewer resources are available to them. In 2016, Woods et al. (28) calculated the CE threshold based on the empirical estimates of opportunity cost, the relationship between a country's GDP per capita, and the value of statistical life. For that reason, they estimated the threshold for different countries with different levels of income. Based on their estimation, Woods et al. suggested the CE threshold to be about 50% of GDP per capita. As appointing the precise CE

threshold level is beyond the scope of our analysis, we interpret and discuss the results with the lowest recommended ICER and leave the decision to policymakers to choose an intervention that best fits their budget. The latest reported GDP per capita of Iran is $4,091. By comparing the results with this threshold, it is shown that all of them, except the low CVD risk group, are highly cost-effective, whereas if the program only targets the people in higher risk groups, it is both cost-saving and improves their health.

Based on the results, the intervention for the low-risk group was not cost-effective as the ICER was undefined. This could be explained by the fact that in this group, just screening is done without offering any intervention. So here, there is a cost for screening without any tangible effect assigned. Since this group does not receive any treatment annually, it is, by all means, sensible that no effects are observed. While it seems to be justifiable, we have every reason to believe that this is the only way to find the groups with a higher CVD risk. Another point that should be mentioned is that the individuals enter and are screened in our model at the age of 40 years. At this age, the proportion of individuals with low CVD risk is significantly higher than that of those with higher CVD risk. The older a person becomes, the more the probability of being in the higher-risk groups will be.

However, the fact that the intervention for the aforementioned group is not cost-effective needs to be approached more comprehensively and conservatively. Screening of this group is the first step and essential for all individuals in the other groups that cannot be disregarded. This means that the other groups can benefit from this, a fact that has not been considered in the analysis of the CE of this group in this study.

From another perspective, a closer observation reveals that screening has a wide range of benefits. For instance, according to the latest national data (29), the prevalence of people with diabetes in Iran is 11.4%, a quarter of whom are undiagnosed (30). This means that at the moment, there are almost 1.5 Iranian people with undiagnosed diabetes. These undiagnosed individuals are discovered only when their complications have started to appear in them. Such complications as retinopathy, nephropathy, and neuropathy are very costly and can impose burdensome pressures on the healthcare system of the country. Other typical examples of this kind are blood hypertension and hyperlipidemia.

It is predicted that huge monetary resources should be allocated to the overall screening of all the individuals, which may not be conveniently supplied. Therefore, appropriate measures could be taken by the authorities to have the costs tailored. This can help manage the financial resources and distribute them as optimally as possible. If the available financial resources do not allow us to screen everyone, it is possible to screen all high-risk people. Although this may not sound optimal, still it has a lot of benefits to offer. In other words, when considered at a higher scale, it can be realized that since the proportion of high CVD risk group individuals outweighs those with lower risks, this may lead to much more favorable results.

It is applicable to have a paper pre-screening. The idea is that alternatively, paper questionnaires can be distributed among both households and health center visitors. These questionnaires aim to detect individuals with higher CVD risks. Typical examples might be those who are obese or have a positive history of CVDs in their intimidate relatives. Once identified, such participants can be invited to health centers to get screened. In this way, we can narrow the target population and screen those who are at higher risk levels.

The study of the literature suggested some good examples of this kind of practice. For example, Chamnan et al. (31), who conducted a modeling study using the data from a prospective cohort study (EPIC-Norfolk), concluded that adopting a stepwise screening approach can prevent the same number of CVD events annually. All the participants of the EPIC-Norfolk study had completed the questionnaire about their lifestyle and drug use and family history of diseases between 1993 and 1997. This population had been observed and followed for 10 years and all CVD events had been recorded. By adopting the Cambridge risk score (a British risk scoring tool) and based on the results of completed questionnaires, the Chamnan group ranked the population CVD risk. Then, they defined and modeled seven different stepwise screening strategies. By comparing the results of a 10-year follow-up of the population with their model, they found that inviting the individuals with a Cambridge risk score >60 might have had the same results as screening the population. According to their report, this strategy could have enabled them to have the same results about the whole population by screening only 60% of the population. Similar results were observed by Móczár and Rurik (32) who concluded that performing screening for a selected target group is most likely to be more cost-effective than screening the whole population. It is essential to consider that although this approach is practical and likely more cost-effective in budget-strained situations, it has some serious ethical issues regarding equity because in this approach, a smoker would get screened and a non-smoker would not. The same is true for a person who has an unhealthy diet or lifestyle. Moreover, while it is conveniently applicable to rural areas, it is not easy to do in urban areas.

Since our literature search revealed no similar studies either in the WHO East Mediterranean region (EMR) or in the Middle East region, inevitably, we compared the results of our analysis with the European studies. Our findings in this analysis, in terms of CE and the trend between the different CVD risks, are similar to the results of Schuetz et al.'s (33) study. The researchers of this study estimated the CE of several different preventive strategies compared with a control scenario in six European countries. By using country-specific data from France, Germany, Denmark, Italy, Poland, and the United Kingdom, they generated six simulated populations of people aged 40–75 years eligible for preventive actions in those countries. Their model showed that the cost per QALY of offering these preventive services to the people in the study cohort ranged from €14,903 for France to €115 for Germany, while it was cost-saving for Poland. Their results showed that the health checks for detecting and managing CVDs at the early stages not only are highly cost-effective but also cost-saving in some scenarios. For example, their analysis for the UK showed that during the 30-year follow-up, the cost per QALY would be €2,426. This ICER in comparison with the UK threshold, which is between 20 and 30 thousand pounds, is highly cost-effective. Moreover, their results demonstrated that the program would be cost-saving if it targets only the top quartile of CVD risk groups. Furthermore, offering prevention checks after the pre-screening of individuals based on some characteristics such as higher age or

obesity would pull the results in the direction of more favorable ICER (33).

Finally, while this screening program can have substantial benefits for individuals with CVD risk, it is essential to consider the potential disutility of lifelong preventive treatments. The search of the literature indicates that the disutility of medications' adverse events (34) and the disutility arising from taking daily medication (35) can play a key role in the decision that leads to non-adherence. It is vital to take into account that some individuals need to be on preventive treatments from their 40s for around 30–40 years. Although the sensitivity analysis demonstrated that this program can likely be cost-effective even with 50% of adherence, it is crucial to enhance treatment compliance through patient education and take effective strategies to increase the engagement of target groups.

Strengths and limitations

Based on our literature search, this analysis is the first CVD risk-based CE to be conducted in Iran, in the Middle East, and the WHO EMR so far. Furthermore, it is one of the few studies which model the individual with and without diabetes separately.

Similar to every CE analysis, this study has several limitations which were mainly caused by a large number of input parameters used in the model. Although the model was designed for the Iranian population, some input parameters were derived from various studies performed in different countries other than Iran directly because of the unavailability or a lack of Iranian data. Furthermore, it was assumed that the intervention effect is equal for all subgroups regardless of their initial CVD risk. Such an assumption can probably generate an underestimation of the intervention effect for high-risk groups and produce an overestimation of the treatment effect for lower CVD risk groups.

In this analysis, the second event of CVDs was not accounted for due to a lack of data. Based on the literature, almost 50% of patients would experience the second or third event during their life once they have had the first event (36). The second event may not necessarily be the same as the first one. For example, a patient who has had a CHD event could have the same event again or can even experience a stroke event. This cannot affect the result of our analysis unfavorably. As the IraPEN intervention causes the first CVD event to decrease or delay, it is logical to assume that the second event in the IraPEN group is lower than the *status quo*. Therefore, we could assume that adding the second CVD event to our model would be in favor of our ICER.

It could be better if we could adopt a lifetime horizon for this analysis. However, as the Iranian statistical data were just available for people aged below 80 years, inevitably 40 years/cycles were employed. The results of Kim et al.'s (37) systematic review, which was done on more than 750 CE analyses, showed that the usage of a lifetime horizon captures all consequences and health benefits most of the time and yields more favorable ICER. Therefore, it is logical to assume that by increasing the time horizon from 40 years to a lifetime, the ICER would be more favorable.

Finally, it should be expected that the effectiveness of the intervention would be lower in the real world than in the model. It could be explained by the fact that the intervention effects, which were used in this study, all had been derived/extracted from a controlled trial setting.

Conclusion

In Iran, CVDs are the leading cause of mortality. Therefore, planning and implementing preventive actions are highly demanded. Our analysis results demonstrated that the IraPEN program implementation is highly cost-effective for all the CVD risk groups, except the low risk without diabetes group, whereas if the program only targets the people in higher risk groups, it is both cost-saving and improves their health.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

AJ contributed to the design, carried out the analysis, and wrote the first draft. RD contributed to the analysis and reviewed the manuscript critically. EA contributed to the design and supervised the process of analysis reviewed the manuscript critically. DK designed the whole project and contributed to the analysis and drafting of the manuscript. All authors read the final manuscript and approved it.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1075277/full#supplementary-material

References

1. Chrysant S. A new paradigm in the treatment of the cardiovascular disease continuum: focus on prevention. *Hippokratia*. (2011) 15:7–11.

2. Azizi F, Hadaegh F, Hosseinpanah F, Mirmiran P, Amouzegar A, Abdi H, et al. Metabolic health in the Middle East and north Africa. *Lancet Diabetes Endocrinol*. (2019) 7:866–79. doi: 10.1016/S2213-8587(19)30179-2

3. Azadnajafabad S, Mohammadi E, Aminorroaya A, Fattahi N, Rezaei S, Haghshenas R, et al. Non-communicable diseases' risk factors in Iran; a review of the present status and action plans. *J Diabetes Metab Disord*. (2021) 22:1–9. doi: 10.1007/s40200-020-00709-8

4. Hadavand Siri F, Khalili D, Hashemi Nazari SS, Ostovar A, Mahdavi A. Adherence to Iran's package of essential noncommunicable diseases (IraPEN) program for regular follow-up to reduce the risk of cardiovascular disease in healthcare centers. *Iran J Endocrinol Metab*. (2020) 22:116–26.

5. World Health Organization. *WHO/ISH Cardiovascular Risk Prediction Charts*. Geneva: World Health Organization (2011). Available online at: https://www.who.int/cardiovascular_diseases/guidelines/Chart_predictions/en/ (accessed February 20, 2019).

6. آمار ملی درگاه . Available online at: https://www.amar.org.ir/ (accessed June 10, 2019) .

7. Naimark DMJ, Kabboul NN, Krahn MD. The half-cycle correction revisited. *Med Decis Making*. (2013) 33:961–70. doi: 10.1177/0272989X13501558

8. Andreson, K, Wilson, P, Odell, P, Kannel, W. An updated coronary risk profile a statement for health professionals. *Circulation*. (1991) 83:356–62. doi: 10.1161/01.CIR.83.1.356

9. Wolf P, D'agostino R, Belanger A, Kannel W. Probability of stroke: a risk profile from the Framingham Study. *Stroke*. (1991) 22:312–8. doi: 10.1161/01.STR.22.3.312

10. Talaei M, Sarrafzadegan N, Sadeghi M, Oveisgharan S, Marshall T, Thomas GN, et al. Incidence of cardiovascular diseases in an Iranian population: the Isfahan Cohort Study. *Arch Iran Med*. (2013) 16:138–44.

11. Grey C, Jackson R, Schmidt M, Ezzati M, Asaria P, Exeter DJ, et al. One in four major ischaemic heart disease events are fatal and 60% are pre-hospital deaths: a national data-linkage study (ANZACS-QI 8). *Eur Heart J*. (2015) 38:172–80. doi: 10.1093/eurheartj/ehv524

12. Shoeibi A, Salehi M, Thrift AG, Kapral MK, Farzadfard MT, Azarpazhooh A, et al. One-year case fatality rate following stroke in the Mashhad Stroke Incidence Study: a population-based study of stroke in Iran. *Int J Stroke*. (2015) 10(Suppl A):96–102. doi: 10.1111/ijs.12611

13. Matza LS, Stewart KD, Gandra SR, Delio PR, Fenster BE, Davies EW, et al. Acute and chronic impact of cardiovascular events on health state utilities. *BMC Health Serv Res*. (2015) 15:173. doi: 10.1186/s12913-015-0772-9

14. Rabani SH, Sardarinia M, Akbarpour S, Azizi F, Khalili D, Hadaegh F. 12-year trends in cardiovascular risk factors (2002-2005 through 2011-2014) in patients with cardiovascular diseases: Tehran lipid and glucose study. *PLoS ONE*. (2018) 13:e0195543. doi: 10.1371/journal.pone.0195543

15. Dagenais GR, Pogue J, Fox K, Simoons ML, Yusuf S. Angiotensin-converting-enzyme inhibitors in stable vascular disease without left ventricular systolic dysfunction or heart failure: a combined analysis of three trials. *Lancet*. (2006) 368:581–8. doi: 10.1016/S0140-6736(06)69201-5

16. Xavier D, Noby M, Pradeep J, Prem P. Letter to the editor. Pattern of drug use in hypertension in a tertiary hospital; a cross sectional study in the inpatients ward. *Indian J Pharmacol*. (2001) 33:456–7.

17. Wright J, Musini V. First-line drugs for hypertension. *São Paulo Med J*. (2010) 128:47. doi: 10.1590/S1516-31802010000100011

18. Wright JM, Musini VM, Gill R. First-line drugs for hypertension. *Cochrane Database Syst Rev*. (2018) 4:CD001841. doi: 10.1002/14651858.CD001841.pub3

19. Law MR, Morris JK, Wald NJ. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *Br Med J*. (2009) 338:b1665. doi: 10.1136/bmj.b1665

20. Brugts J, Yetgin T, Hoeks S, Gotto A, Shepherd J, Westendorp R, et al. The benefits of statins in people without established cardiovascular disease but with cardiovascular risk factors: meta-analysis of randomised controlled trials. *BMJ*. (2009) 338:b2376. doi: 10.1136/bmj.b2376

21. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HAW. 10-year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med*. (2008) 359:1577–89. doi: 10.1056/NEJMoa0806470

22. Lanier J, Bury D, Richardson S. Diet and physical activity for cardiovascular disease prevention. *Am Fam Physician*. (2016) 93:919–24.

23. شعار سال وزارت بهداشت، درمان و آموزش پزشکی. Available online at: http://behdasht.gov.ir/ (accessed March 29, 2019).

24. Javanbakht M, Jamshidi AR, Baradaran HR. Estimation and prediction of avoidable health care costs of cardiovascular diseases and type 2 diabetes through adequate dairy food consumption: a systematic review and micro simulation modeling study. *Arch Iran Med*. (2018) 21:213–22.

25. *GDP per Capita (current US$) - Iran, Islamic rep. Data*. Available online at: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=IR (accessed January 21, 2023).

26. Marseille E, Larson B, Kazi DS, Kahn JG, Rosen S. Thresholds for the cost–effectiveness of interventions: alternative approaches. *Bull World Health Organ*. (2014) 93:118–24. doi: 10.2471/BLT.14.138206

27. Revill P, Walker SM, Madan J, Ciaranello A, Mwase T, Gibb DM et al. *Using Cost-effectiveness Thresholds to Determine Value for Money in Low-and Middle-income Country Healthcare Systems: Are Current International Norms fit For Purpose?* York, UK: Centre for Health Economics, University of York (2014). (CHE Research Paper; 98).

28. Woods B, Revill P, Sculpher M, Claxton K. Country-Level cost-effectiveness thresholds: initial estimates and the need for further research. *Value Health*. (2016) 19:929–35. doi: 10.1016/j.jval.2016.02.017

29. Esteghamati A, Etemad K, Koohpayehzadeh J, Abbasi M, Meysamie A, Noshad S, et al. Trends in the prevalence of diabetes and impaired fasting glucose in association with obesity in Iran: 2005–2011. *Diabetes Res Clin Pract*. (2014) 103:319–27. doi: 10.1016/j.diabres.2013.12.034

30. Noshad S, Afarideh M, Heidari B, Mechanick JI, Esteghamati A. Diabetes care in Iran: where we stand and where we are headed. *Ann Glob Health*. (2016) 81:839. doi: 10.1016/j.aogh.2015.10.003

31. Chamnan P, Simmons RK, Khaw KT, Wareham NJ, Griffin SJ. Estimating the population impact of screening strategies for identifying and treating people at high risk of cardiovascular disease: modelling study. *BMJ*. (2010) 340(apr23 2):c1693. doi: 10.1136/bmj.c1693

32. Móczár C, Rurik I. Comparison of cardiovascular risk screening methods and mortality data among Hungarian primary care population: preliminary results of the first government-financed managed care program. *Slovenian J Public Health*. (2015) 54:154–60. doi: 10.1515/sjph-2015-0022

33. Schuetz CA, Alperin P, Guda S, Herick AV, Cariou B, Eddy D, et al. A standardized vascular disease health check in Europe: a cost-effectiveness analysis. *PLoS ONE*. (2013) 8:e0066454. doi: 10.1371/journal.pone.0066454

34. Epstein D, García-Mochón L, Kaptoge S, Thompson SG. Modeling the costs and long-term health benefits of screening the general population for risks of cardiovascular disease: a review of methods used in the literature. *Eur J Health Econ*. (2015) 17:1041–53. doi: 10.1007/s10198-015-0753-2

35. Cutler RL. Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. *BMJ Open*. (2018) 8:e016982. doi: 10.1136/bmjopen-2017-016982

36. Bansilal S, Castellano JM, Fuster V. Global burden of CVD: focus on secondary prevention of cardiovascular disease. *Int J Cardiol*. (2015) 1:S1–7. doi: 10.1016/S0167-5273(15)31026-3

37. Kim DD, Wilkinson CL, Pope EF, Chambers JD, Cohen JT, Neumann PJ. The influence of time horizon on results of cost-effectiveness analyses. *Expert Rev Pharmacoecon Outcomes Res*. (2017) 17: 615–23. doi: 10.1080/14737167.2017.1331432

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 18 April 2023
DOI 10.3389/fpubh.2023.1123552

OPEN ACCESS

EDITED BY
Ke Yan,
National University of Singapore, Singapore

REVIEWED BY
Muhammad Imran Khan,
Hanyang University, Republic of Korea
Arinjita Bhattacharyya,
University of Louisville, United States

*CORRESPONDENCE
Xiaoyu Xi
✉ xixy@cpu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 14 December 2022
ACCEPTED 03 April 2023
PUBLISHED 18 April 2023

CITATION
Wan C, Wang Q, Xu Z, Huang Y and Xi X (2023) Mapping health assessment questionnaire disability index onto EQ-5D-5L in China. *Front. Public Health* 11:1123552. doi: 10.3389/fpubh.2023.1123552

Mapping health assessment questionnaire disability index onto EQ-5D-5L in China

Chuchuan Wan†, Qiqi Wang†, Zhaoqi Xu, Yuankai Huang and Xiaoyu Xi*

The Research Center of National Drug Policy & Ecosystem, China Pharmaceutical University, Nanjing, China

Objective: This research aimed to develop the more accurate mapping algorithms from health assessment questionnaire disability index (HAQ-DI) onto EQ-5D-5L based on Chinese Rheumatoid Arthritis patients.

Methods: The cross-sectional data of Chinese RA patients from 8 tertiary hospitals across four provincial capitals was used for constructing the mapping algorithms. Direct mapping using Ordinary least squares regression (OLS), the general linear regression model (GLM), MM-estimator model (MM), Tobit regression model (Tobit), Beta regression model (Beta) and the adjusted limited dependent variable mixture model (ALDVMM) and response mapping using Multivariate Ordered Probit regression model (MV-Probit) were carried out. HAQ-DI score, age, gender, BMI, DAS28-ESR and PtAAP were included as the explanatory variables. The bootstrap was used for validation of mapping algorithms. The average ranking of mean absolute error (MAE), root mean square error (RMSE), adjusted R^2 (adjR^2) and concordance correlation coefficient (CCC) were used to assess the predictive ability of the mapping algorithms.

Results: According to the average ranking of MAE, RMSE, adjR^2, and CCC, the mapping algorithm based on Beta performed the best. The mapping algorithm would perform better as the number of variables increasing.

Conclusion: The mapping algorithms provided in this research can help researchers to obtain the health utility values more accurately. Researchers can choose the mapping algorithms under different combinations of variables based on the actual data.

KEYWORDS

HAQ-DI, EQ-5D-5L, mapping, Rheumatoid Arthritis, health utility, China

Introduction

Rheumatoid Arthritis (RA) is a chronic autoimmune disorder with symmetrical, recurrent, incurable and highly disabling (1, 2). RA has negative impact on patients physically, psychologically, and socially, such as leading to reduced daily activities, depression, changes in career plans, and reduced financial income, etc., which seriously affects the quality of life for patients (3). RA also imposes a huge financial burden on patients and society. Globally, RA imposes the greatest burden of all rheumatic diseases (4). In mainland China, the annual economic burden of RA patients is as high as 72 million Chinese Yuan (CNY). Considering the influence of per capita disability adjusted life years (DALYs), the annual economic burden is as high as 902 million CNY and annual economic burden per capita is 15,718 CNY (5).

Given the severe burden of RA and finite health resources, it is necessary to assess the value of interventions for RA. Cost-utility analysis (CUA) is the most widely used economic evaluation method (6), in which quality adjusted life years (QALYs) is adopted as the main health outcome (7). As an essential metric for calculating QALY, health utility value (HUV) is often obtained through a preference-based measure (PBM), like EuroQol Five-dimensions Questionnaire (EQ-5D), Short Form Six-dimension (SF-6D), Health Utilities Index (HUI) (8–10). As the most commonly used PBM, EQ-5D includes EQ-5D-3L, EQ-5D-5L, and EQ-5D-Y, with EQ-5D-5L being more widely used in China. In clinical, the health assessment questionnaire disability index (HAQ-DI) is widely used to assess the quality of life and function of RA patients (11). However, researchers can't get HUV of patient from HAQ-DI which is a non-preference-based measure.

Up to now, a large number of researches have demonstrated that the value of HAQ-DI can be converted into HUV through mapping (12–21). In Most researches, EQ-5D-3L was chosen as the target scale and the HAQ-DI total score as a single independent variable. Part of researches also would include sociodemographic or clinical characteristics as covariates. Interestingly, some of the CUA conducted on Chinese RA patients chose to use the mapping algorithm developed on Spanish populations (19, 22–24), although the mapping algorithms based on Chinese populations were available (13). Now, two researches have developed the mapping algorithms from HAQ-DI to EQ-5D-5L based on Chinese population. But they both suffered from single sample source, few models and variables selected.

Hence, we aimed to construct a mapping algorithm from HAQ-DI to EQ-5D-5L based on richer data sources, model selection and variable including. And the algorithm can be used to address the lack of HUV in health technology assessment for Chinese RA patients.

Materials and methods

We adopted the cross-sectional data of RA patients, in which the EQ-5D-5L and HAQ-DI were used to measure and value HRQOL. The data were collected by trained investigators based on quota sampling during June to July 2020, which from 8 tertiary hospitals across four provincial capitals Nanjing, Hangzhou, Chengdu, and Shijiazhuang (two of each). This study complied with the principles of the Declaration of Helsinki and approved by the Clinical Trial Ethics Committee of Huashan Hospital Affiliated to Fudan University (Reference Number 2019-252).

Sample

Given the available research resources and rules of thumb, 25 patients were recruited for each center, a total of 200 patients. The inclusion criteria were: (1) Informed and voluntary; (2) 18 to 70 years; (3) Diagnosed with RA according to the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification diagnostic criteria (score $\geq$6) (25). The exclusion criteria were: (1) non-Chinese; (2) Gravidae and the patients who are unconscious; (3) Suffering from other serious diseases that seriously affect the quality of life, like tumors, myocardial infarct.

Data collection

Data collection for each center was done by 2 interviewers who were systematically trained and knowledgeable about the content and methodology of research. The interviewers introduced the research to patients and doctors in the corresponding departments with the permission of the hospital administrator. For patients and their attending doctors willing to participate in the research, the interviewers would provide written informed consent to them. Then, the patients and doctors would be provided with an electronic device containing the research questionnaire and asked to complete the respective questionnaires independently in a quiet room without any guidance from the interviewers. The questionnaires completed were reviewed by the interviewers and uploaded to the auditors if no obvious errors or blanks. And the data would be digitalized and reviewed by 2 auditors.

Questionaries

Questionnaires were designed for doctors and patients separately. The literature (26, 27) and experts' opinions were drawn upon. According to the results of a pilot survey conducted in 2 tertiary hospitals in Nanjing, we revised and formed the final questionnaires. The rationality, readability and comprehensibility for questionnaires had been affirmed by the experts and supported by the pilot survey.

The questionnaire had two parts for patients. Part 1 collected socio-demographic information, including age, gender, height, weight, region, education level. Part 2 collected some health status indicators which reported by patients through EQ-5D-5L、HAQ-DI、the patient's assessment of arthritis pain visual analog scale (PtAAP-VAS) and the patient's global assessment of disease activity visual analog scale (PtGADA-VAS). PtAAP-VAS and PtGADA-VAS were used to assesses arthritis pain and disease activity, of which 0 means no symptoms and 100 means severe symptoms.

The questionnaire also had two parts for doctors. Part 1 collected clinical information for patients, including high-sensitivity C-reactive protein (CRP) (unit: mg/L), erythrocyte sedimentation rate (ESR) (unit: mm/h), swollen joints count (SJC) and tender joints count (TJC). Part 2 collected the physician's assessment of the patients' disease activity through the Physician's global assessment of disease activity visual analog scale (PhGADA-VAS), of which 0 means no symptoms and 100 means severe symptoms. In addition, we calculated the 28 joint counts (DAS28) scores (28), including DAS28-CRP score and DAS28-ESR score, based on CRP, ESR SJC and TJC. Disease activity is divided into four states, including remission (DAS28 scores $<$2.6), low activity group (2.6$<$ DAS28 scores $<$3.2), moderate activity group (3.2$\leq$ DAS28 scores $<$5.1), and high activity group (DAS28 scores $\geq$5.1) (28).

EQ-5D-5L

The EQ-5D-5L is more sensitive compared with EQ-5D-3L and has been widely used to measure HUV. The EQ-5D-5L essentially consists of 2 parts: the EQ-5D descriptive system and the EQ visual analog scale (EQ VAS) (29). The descriptive system contains five dimensions, including three functional dimensions (mobility, self-care, and usual activities) and two somatic symptom dimensions (pain/discomfort, and anxiety/depression), each of which is divided into five levels (no problems, slight problems, moderate problems, severe problems, and unable to/extreme problems) and produces a total of 3125 (5^5) health states (29, 30). The EQ VAS assesses the self-reported health status of subjects through a straight line (0: the worst health you can imagine; 100: the best health you can imagine). The reliability and validity of EQ-5D-5L have been verified in China (31). The EQ-5D-5L utility scores in this study were calculated using the China value set (32).

HAQ-DI

The HAQ has two versions, the full HAQ and the short HAQ. The short HAQ which is used frequently contains the HAQ-DI, the VAS Pain Scale, and the VAS Patient Global. Further, the HAQ-DI which was developed by Stanford University in 1978 is often used by itself, particularly but not exclusively in the rheumatic disease (33). The HAQ-DI assesses a patient's level of functional ability with 20 questions in 8 categories, including dressing, rising, eating, walking, hygiene, reach, grip, and usual activities. Each category consists of 2 or 3 items and each item contains 4 level (0 means no difficulty, 1 means some difficulty, 2 means much disability, and 3 means unable to do). The score for each category is the highest score for the item in this category and the overall HAQ-DI score is the average of 8 categories, within 0 to 3 (33, 34).

Data analysis

Descriptive statistics

Descriptive statistics (mean and standard deviation (SD) for continuous variables, frequency and percentage for categorical variables) were used for the sample characteristics. The distributions of EQ-5D-5L utility score and HAQ-DI score were showed through the figures.

Correlation test

The estimation of the mapping algorithm requires conceptual overlap between the source scale and the target scale (35–38). Spearman rank correlations were used to test the correlations between the HAQ-DI scores and EQ-5D-5L utility scores, so that we could ensure the degree of conceptual overlap. In addition, the correlations of different variables were tested by Spearman rank correlations to guarantee low collinearity among the independent variables included in the mapping algorithm. The strength of correlation could be divided to 4 level (very weak = 0–0.19; weak = 0.20–0.39; moderate = 0.40–0.59; strong = 0.60–0.79; and very strong = 0.80–1.00) (39, 40).

Mapping model

Mapping consists of two broad approaches, direct mapping and response mapping. We used seven statistical models for developing a simpler and more accurate mapping algorithm, based on guidelines and previous researches (35). Ordinary least squares regression (OLS), the general linear regression model (GLM), MM-estimator model (MM), Tobit regression model (Tobit), Beta regression model (Beta) and the adjusted limited dependent variable mixture model (ALDVMM) were used for direct mapping and Multivariate Ordered Probit regression model (MV-Probit) for response mapping.

OLS, which assumes a linear relationship between the dependent variables and independent variables, is the most commonly used in direct mapping due to its simplicity (40, 41). But OLS performs poorly for predicting poor or full health, and the predicted values may be outside of the reasonable range (18, 36, 42). GLM, a flexible form of OLS, allows the outcome variables to have non-normal error distributions (43, 44). As one of robust regression technique, MM builds on minimizing some function of the residuals and a measure of dispersion of the residuals to achieve a high breakdown point with high efficiency (45–47). Given the EQ-5D-5L score has a ceiling effect (36, 43), and our measurements with a left-skewed distribution are clustered at 1, Tobit was used for mapping. Tobit can estimate the linear relationships among variables, as a censored model. But it is sensitive to violations of heteroscedasticity or non-normality (48, 49). Beta can avoid the predicted value fall outside of the reasonable range by setting the value of the dependent variable between 0-1. In addition, Beta is also suitable for heteroscedasticity or non-normality (50, 51). The following formula was used to adjust EQ-5D-5L scores to range of 0 to 1: *adjusted score = (original score+0.391)/1.391* (The range of EQ-5D-5L score is −0.391 to 1, based on the Chinese value set. If the original score was −0.391 or 1, the adjusted score was added or subtracted e^{-12} to ensure it fell between 0 and 1.). ALDVMM was developed as a mixture model for dealing with the distributional features of EQ-5D-3L. It could effectively capture the multimodal properties of EQ-5D score, boundary value and the gap between the value for full health and other health states (18). ALDVMM has been used in numerous previous researches (17, 52, 53), and confirmed is applicable to EQ-5D-5L (36, 52, 54–56). In response mapping, the dimension results obtained by the mapping algorithm are used to calculate the health utility values, based on the value set (36). Order Probit and Order Logit, etc., are commonly used models, but they are unable to explain the correlations between different dimensions. Multivariate ordered Probit method(MV-Probit) developed by Conigliani can solve the problem well (57).

Variables

We chose the EQ-5D-5L total score and values of each dimension of EQ-5D-5L as the dependent variable for direct and response mapping, respectively. The HAQ-DI total score was included as the main independent variable for all mappings. The

TABLE 1 Combinations of variables.

Combination	Explanatory variables
Combination 1	HAQ-DI
Combination 2	HAQ-DI, PtAAP
Combination 3	HAQ-DI, Age, Gender
Combination 4	HAQ-DI, Age, BMI, Gender
Combination 5	HAQ-DI, DAS28-ESR, PtAAP, Age, BMI, Gender

HAQ-DI, the health assessment questionnaire disability index; PtAAP, the patient's assessment of arthritis pain; BMI, body mass index; DAS28, 28 joint counts; ESR, erythrocyte sedimentation rate.

reason for we didn't chose the scores of different parts of HAQ-DI as independent variables for response mapping was that the sample size was slightly lacking. To obtain a more accurate mapping algorithm, we also included factors such as age, gender, BMI, DAS28-ESR, and PtAAP as independent variables based on the correlation between the EQ-5D-5L utility scores and each variable, as well as between the variables. It should be noted that we had scaled PtAAP (divided by 100) for calculation and presentation in the regression. And different combinations of variables were set in order to balance the accuracy, simplicity and generalizability of the mapping algorithms (Table 1).

Validation and comparison of mapping algorithms

Given our sample was not rich, the bootstrap was used for validation of mapping algorithms (13). Firstly, a bootstrap sample of the same size as the original sample was drawn from the original sample. Secondly, mapping algorithms were fitted from the bootstrap sample using all statistics models, for all combinations of variables. Thirdly, the health utility values predicted using the mapping algorithms were compared with the observed health utility values, in the original sample. Finally, the mean of the health utility score predicted (MEAN-P), absolute difference between mean predicted and mean observed scores (ADM), mean absolute error (MAE), root mean square error (RMSE), adjusted R^2 (adjR^2) and concordance correlation coefficient (CCC) were calculated, recorded (43). After repeating the aforementioned four steps 500 times, the ranking of four indictors (MAE, RMSE, adjR^2, CCC) were recorded and averaged for each combination. The mapping algorithm with the best average ranking was the optimal mapping algorithm in different combination of variables.

All statistical analyses were performed by stata15, programs R and Microsoft® Excel 2019.

Results

Sociodemographic characteristics and patient-reported outcomes

A total of 172 eligible patients were enrolled in the researches (Table 2). Their mean age (SD) was 50.82 (12.09) years, which was not significantly different from that of the general Chinese RA patient population ($p > 0.1$) (2, 26, 58). The proportion of female was 63.37%. A total of 126 health states were obtained and the mean scores (SD) were 0.59 (0.28) and 1.49 (0.60) for EQ-5D-5L utility score and HAQ-DI score, respectively. The distributions were left skewed and normally for EQ-5D-5L utility scores and HAQ-DI scores, respectively (Figure 1).

Correlation test results

Correlations between EQ-5D-5L utility score, HAQ-DI score, age, gender, BMI and DAS28-ESR, etc. were provided in Appendix 1 (see Supplementary material). There was a strong negative correlation between EQ-5D-5L utility score and HAQ-DI score (−0.7067). But other variables such as age, BMI and DAS28-ESR, etc. showed weak or moderate correlation with EQ-5D-5L utility score.

Mapping algorithm performance

The results of mapping algorithms performance were presented in Table 3. A total of 35 mapping algorithms were fitted based on 5 combinations of variables and 7 statistics models. Limited the sample size and the convergence of the model, only the ALDVMM with one component was estimated.

The MEAN-P ranged from 0.5808 (ALDVMM of combination 5) to 0.6124 (MM of combination 3), where the OLS of combination 2 had the closest predicted score (ADM = 0.0002) to the observed mean score. The mapping algorithms of combination 5 had the lower MAE and RMSE and higher adjR^2 and CCC as the number of variables increasing. Of all mapping algorithms, the range was 0.1263 (Beta of combination 5) to 0.1668 (GLM of combination 1) for MAE, 0.1644 (Beta of combination 5) to 0.2092 (GLM of combination 1) for RMSE, 0.4394 (GLM of combination 1) to 0.6432 (Beta of combination 5) for adjR^2, 0.6016 (GLM of combination 1) to 0.8034 (Beta of combination 5) for CCC.

Best performing mapping algorithm

The average group ranking (AGR) showed that Beta performs the best, followed by MVROD-Probit. It should be note that the AGR of MV-Probit and Beta both are 1 for combination 3 and 4. The Beta performed better in MAE and CCC, but worse in RMSE and adjR^2 than MVROD-Probit. In addition, compared with Beta, MVROD-Probit performed better in ADM for each combination. But Beta performed the best in MAE, RMSE, adjR^2 and CCC for combination 1, 2, and 5. Based on these results, we thought that the best performing mapping algorithm was the one constructed based on the Beta, for each combination. As for the best combination, the results showed that the more variables incorporated, the better the mapping algorithm performed. But we did not think the combination 5 was the best because it required more variables. The reality was that we may could not obtain

10.3389/fpubh.2023.1123552

TABLE 2 Socio-demographic characteristics and patient-reported outcomes.

Characteristics ($N = 172$)	Mean ± SD/N (%)	Median	Min	Max
Age (years)	50.82 ± 12.09	53	20	70
Gender				
Male	63.00 (36.63%)			
Female	109.00 (63.37%)			
Region				
Urban	78.00 (45.35%)			
Rural	94.00 (54.65%)			
Education				
Below primary school	40.00 (23.26%)			
Primary school	35.00 (20.35%)			
junior middle school	27.00 (15.70%)			
High school/technical secondary school	40.00 (23.26%)			
Undergraduate/Junior college	21.00 (12.21%)			
Master degree or above	9.00 (5.23%)			
BMI	22.75 ± 3.72	22.30	14.64	35.67
EQ-5D-5L	0.59 ± 0.28	0.66	−0.19	1.00
EQ VAS	47.59 ± 19.85	48.20	10.00	90.00
PtAAP	63.85 ± 18.06	65.00	7.00	100.00
HAQ-DI	1.49 ± 0.60	1.50	0.25	2.88
DAS28-CRP	5.45 ± 1.20	5.45	2.01	8.32
DAS28-ESR	5.56 ± 1.37	5.66	1.77	9.09
SJC	14.04 ± 9.04	12.00	0.00	54.00
TJC	22.76 ± 14.44	19.00	0.00	68.00
ESR	48.63 ± 29.22	40.00	2.00	121.00
CRP	30.53 ± 36.31	13.00	0.50	177.10

SD, standard deviation; BMI, body mass index; EQ VAS, EuroQol visual analog scale; PtAAP, the patient's assessment of arthritis pain; HAQ-DI, the health assessment questionnaire disability index; DAS28, 28 joint counts; SJC, swollen joints count; TJC, tender joints count; ESR, erythrocyte sedimentation rate; CRP, high-sensitivity C-reactive protein.

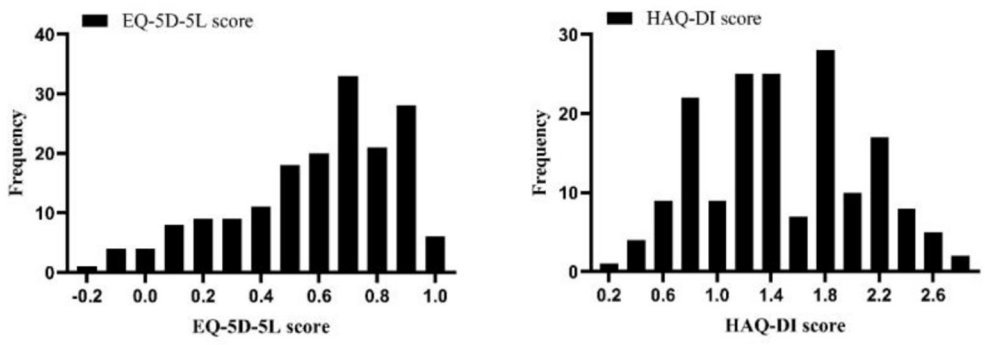


FIGURE 1
Distributions of EQ-5D-5L utility scores and HAQ-DI scores.

the data for these variables. Thus, we recommended choosing the combination according to the actual available data and selecting the mapping algorithm constructing based on Beta. Figure 2 demonstrated the consistency between the observed EQ-5D-5L utility score and predicted EQ-5D-5L utility score, based on Beta, for each combination. The Pearson correlation coefficients were

TABLE 3 The results of mapping algorithms performance.

Combination	Model	MEAN-P (SD)	ADM	MAE (SD)	GR (MAE)	RMSE (SD)	GR (RMSE)	adjR^2 (SD)	GR (adjR^2)	CCC (SD)	GR (CCC)	AGR
Combination 1	OLS	0.5902 (0.0151)	0.0004	0.1552 (0.0025)	5	0.1994 (0.0011)	4	0.4907 (0.0060)	4	0.6640 (0.0169)	4	3
	GLM	0.5947 (0.0153)	0.0041	0.1668 (0.0026)	7	0.2092 (0.0011)	7	0.4396 (0.0070)	7	0.6017 (0.0181)	7	7
	MM	0.6060 (0.0169)	0.0155	0.1535 (0.0018)	3	0.2002 (0.0020)	6	0.4869 (0.0109)	6	0.6631 (0.0186)	5	6
	ALDVMM	0.5838 (0.0154)	0.0068	0.1553 (0.0033)	6	0.1981 (0.0015)	3	0.4972 (0.0078)	3	0.6579 (0.0191)	6	5
	Tobit	0.5925 (0.0152)	0.0019	0.1545 (0.0025)	4	0.1995 (0.0011)	5	0.4904 (0.0061)	5	0.6684 (0.0162)	3	3
	Beta	**0.5936 (0.0148)**	**0.0031**	**0.1497 (0.0024)**	**1**	**0.1951 (0.0014)**	**1**	**0.5128 (0.0071)**	**1**	**0.6937 (0.0153)**	**1**	**1**
	MV-Probit	0.5892 (0.0149)	0.0013	0.1511 (0.0030)	2	0.1954 (0.0012)	2	0.5110 (0.0064)	2	0.6766 (0.0172)	2	2
Combination 2	OLS	0.5904 (0.0150)	0.0002	0.1512 (0.0020)	6	0.1949 (0.0015)	4	0.5108 (0.0078)	4	0.6847 (0.0174)	5	5
	GLM	0.5953 (0.0149)	0.0047	0.1640 (0.0022)	7	0.2065 (0.0012)	7	0.4509 (0.0100)	7	0.6152 (0.0177)	7	7
	MM	0.6021 (0.0169)	0.0115	0.1505 (0.0018)	4	0.1957 (0.0025)	6	0.5065 (0.0130)	6	0.6878 (0.0196)	4	6
	ALDVMM	0.5844 (0.0146)	0.0062	0.1500 (0.0028)	3	0.1928 (0.0017)	3	0.5213 (0.0089)	3	0.6839 (0.0188)	6	3
	Tobit	0.5925 (0.0151)	0.0019	0.1505 (0.0020)	5	0.1949 (0.0015)	5	0.5107 (0.0078)	5	0.6887 (0.0169)	3	4
	Beta	**0.5952 (0.0141)**	**0.0047**	**0.1454 (0.0018)**	**1**	**0.1889 (0.0020)**	**1**	**0.5403 (0.0099)**	**1**	**0.7179 (0.0158)**	**1**	**1**
	MV-Probit	0.5887 (0.0146)	0.0018	0.1474 (0.0024)	2	0.1900 (0.0019)	2	0.5349 (0.0097)	2	0.6987 (0.0185)	2	2
Combination 3	OLS	0.5900 (0.0139)	0.0006	0.1455 (0.0030)	6	0.1829 (0.0013)	5	0.5666 (0.0064)	5	0.7328 (0.0121)	4	5
	GLM	0.5957 (0.0144)	0.0052	0.1586 (0.0029)	7	0.1966 (0.0015)	7	0.4993 (0.0131)	7	0.6637 (0.0151)	7	7
	MM	0.6125 (0.0169)	0.0219	0.1424 (0.0023)	3	0.1846 (0.0027)	6	0.5582 (0.0135)	6	0.7291 (0.0157)	6	6
	ALDVMM	0.5842 (0.0138)	0.0063	0.1450 (0.0033)	5	0.1811 (0.0017)	3	0.5749 (0.0083)	3	0.7309 (0.0131)	5	4
	Tobit	0.5912 (0.0140)	0.0007	0.1444 (0.0031)	4	0.1823 (0.0014)	4	0.5692 (0.0068)	4	0.7369 (0.0123)	3	3
	Beta	**0.5935 (0.0133)**	**0.0029**	**0.1390 (0.0033)**	**1**	**0.1777 (0.0016)**	**2**	**0.5909 (0.0078)**	**2**	**0.7617 (0.0101)**	**1**	**1**
	MV-Probit	0.5903 (0.0134)	0.0003	0.1406 (0.0035)	2	0.1777 (0.0015)	1	0.5909 (0.0070)	1	0.7506 (0.0114)	2	1
Combination 4	OLS	0.5897 (0.0134)	0.0009	0.1435 (0.0025)	5	0.1782 (0.0016)	5	0.5859 (0.0075)	5	0.7502 (0.0113)	4	5
	GLM	0.5955 (0.0139)	0.0050	0.1562 (0.0025)	7	0.1921 (0.0016)	7	0.5190 (0.0160)	7	0.6849 (0.0131)	7	7
	MM	0.6102 (0.0173)	0.0196	0.1413 (0.0022)	3	0.1803 (0.0036)	6	0.5763 (0.0180)	6	0.7451 (0.0167)	5	6
	ALDVMM	0.5836 (0.0143)	0.0069	0.1436 (0.0036)	6	0.1775 (0.0030)	3	0.5894 (0.0143)	3	0.7449 (0.0165)	6	4
	Tobit	0.5903 (0.0138)	0.0003	0.1424 (0.0027)	4	0.1777 (0.0018)	4	0.5882 (0.0088)	4	0.7544 (0.0120)	3	3

(Continued)

10.3389/fpubh.2023.1123552

TABLE 3 (Continued)

Combination	Model	MEAN-P (SD)	ADM	MAE (SD)	GR (MAE)	RMSE (SD)	GR (RMSE)	adjR^2 (SD)	GR (adjR^2)	CCC (SD)	GR (CCC)	AGR
	Beta	**0.5936 (0.0129)**	**0.0031**	**0.1372 (0.0030)**	**1**	**0.1739 (0.0021)**	**2**	**0.6058 (0.0100)**	**2**	**0.7748 (0.0092)**	**1**	**1**
	MV-Probit	0.5902 (0.0129)	0.0004	0.1387 (0.0030)	2	0.1736 (0.0017)	1	0.6071 (0.0077)	1	0.7655 (0.0104)	2	1
Combination 5	OLS	0.5898 (0.0131)	0.0008	0.1360 (0.0027)	5	0.1709 (0.0019)	4	0.6147 (0.0088)	4	0.7756 (0.0110)	4	4
	GLM	0.5963 (0.0136)	0.0057	0.1539 (0.0026)	7	0.1888 (0.0021)	7	0.5295 (0.0240)	7	0.6993 (0.0129)	7	7
	MM	0.6110 (0.0183)	0.0204	0.1342 (0.0029)	3	0.1737 (0.0049)	5	0.6015 (0.0234)	5	0.7735 (0.0156)	5	5
	ALDVMM	0.5809 (0.0160)	0.0097	0.1434 (0.0081)	6	0.1784 (0.0090)	6	0.5791 (0.0438)	6	0.7401 (0.0372)	6	6
	Tobit	0.5888 (0.0134)	0.0018	0.1353 (0.0031)	4	0.1703 (0.0020)	3	0.6173 (0.0093)	3	0.7817 (0.0127)	3	3
	Beta	**0.5956 (0.0123)**	**0.0051**	**0.1264 (0.0034)**	**1**	**0.1644 (0.0028)**	**1**	**0.6433 (0.0123)**	**1**	**0.8034 (0.0086)**	**1**	**1**
	MV-Probit	0.5893 (0.0124)	0.0012	0.1283 (0.0036)	2	0.1645 (0.0021)	2	0.6429 (0.0092)	2	0.7944 (0.0102)	2	2

OLS, Ordinary least squares regression; GLM, the general linear regression model; MM, MM-estimator model; Tobit, Tobit regression model; Beta, Beta regression model; ALDVMM, the adjusted limited dependent variable mixture model; MV-Probit, Multivariate Ordered Probit regression model; MEAN-P, mean of the health utility score predicted; ADM, absolute difference between mean predicted and mean observed scores; GR, group ranking; MAE, mean absolute error; RMSE, root mean square error; adjR^2, adjusted R^2; CCC, concordance correlation coefficient; AGR, average group ranking. The best performers have been bolded.

0.6228, 0.6411, 0.6889, 0.6975, and 0.7314 for combination 1 to 5, respectively. They indicated a high correlation between the observed EQ-5D-5L score and predicted EQ-5D-5L score.

The mapping algorithm parameters

Table 4 presents the parameters of mapping algorithms constructed based on Beta for predicting EQ-5D-5L utility scores from HAQ-DI scores. For all combinations, HAQ-DI score, gender, BMI and PtAAP were significant predictors of EQ-5D-5L utility scores, but age and DAS28-ESR were insignificant predictors. According to the formula (see Methods) for adjusting EQ-5D-5L scores, the mapping algorithms formula can be showed as following:

$$EQ-5D-5L\ score = 1.391 \times \frac{e^{constant+\sum_1^i \beta_i X_i}}{1+e^{constant+\sum_1^i \beta_i X_i}} - 0.391$$

[i is the count of independent variables included, β_i is the coefficient (parameters showed in Table 4), X_i is the independent variable, such as HAQ-DI, gender, age, etc.].

Discussion

This research developed the mapping algorithms for estimating EQ-5D-5L utility scores from HAQ-DI scores in Chinese RA patients. Up to now, two researches have constructed the mapping algorithm from HAQ-DI to EQ-5D-5L based on Chinese RA patients, of which Thoma's research (13) used the Beta and MV-Probit and Dexin ZHOU's research (12) used OLS and Tobit. But they used different statistical models and, respectively concluded that OLS (Dexin ZHOU) and MVROD-Probit (Thomas) were the best models for developing the mapping algorithms. This was difficult for researchers to choose the optimal mapping algorithm to conduct relative researches. Furthermore, ALDVMM was confirmed by several studies to have significant advantages in predicting EQ-5D-5L utility scores (36, 55, 59). Ducournau P indicated that the relationship between HAQ-DI scores and EQ-5D utility scores was non-linear in his study (21). Based on these results, this research incorporated additional models to construct mapping algorithms, including OLS, GLM, MM, ALDVMM, Tobit, Beta, and MVROD-Probit. Compared with the single sample source of previous two researches, this research increased to four sample sources, although also not particularly rich in sample size. In addition, the variables included in this research were richer than previous researches which only includes HAQ-DI score, Pain VAS (equivalent to PtAAP), EQ-VAS and PrGA (equivalent to PhGADA). The mapping algorithms in this research were constructed after considering linear model, mixed model, response mapping and a large number of variables. It was benefit for obtaining the health utility values of Chinese RA patients, and then conducting pharmacoeconomic evaluation to enhance the efficiency of healthcare resource allocation.

The selection of variables referred to the existing literature and the correlations test results (Appendix 1). Among the previous

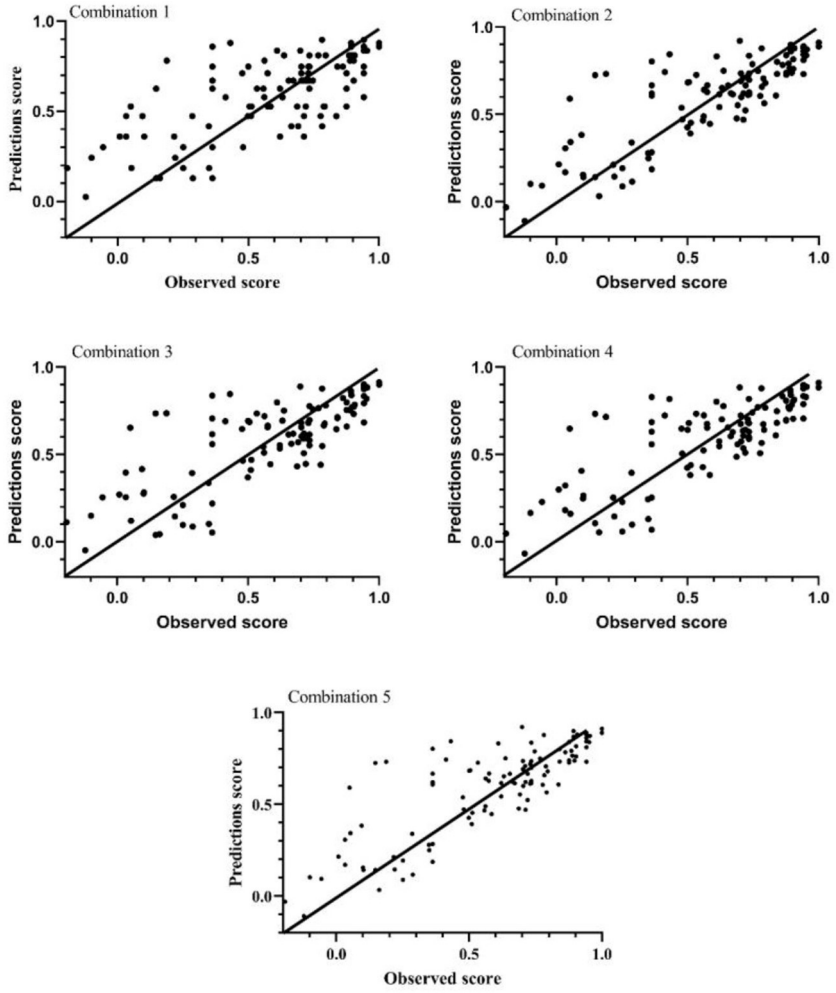


FIGURE 2
The observed and predicted EQ-5D-5L utility score.

TABLE 4 The parameters of mapping algorithm from EQ-5D-5L to HAQ-DI based on beta.

	Combination 1	Combination 2	Combination 3	Combination 4	Combination 5
Constant	3.02428***	3.48797***	3.68661***	4.35968***	**4.93095*** **
HAQ-DI	−1.34965***	−1.26531***	−1.32507***	−1.32570***	**-1.22172*** **
Gender			−0.68938***	−0.71406***	**−0.76041*** **
Age			−0.00473	−0.00300	**−0.00137**
BMI				−0.03246*	**−0.02798***
PtAAP		−0.90418**			**−1.11943*** **
DAS28-ESR					**−0.02660**

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
The PtAAP scores used here should be the scaled scores (divided by 100).
HAQ-DI, the health assessment questionnaire disability index; PtAAP, the patient's assessment of arthritis pain; BMI, body mass index; DAS28, 28 joint counts; ESR, erythrocyte sedimentation rate. The best performers have been bolded.

researches that develop the mapping algorithms from HAQ-DI to EQ-5D-5L, researchers often included HAQ-DI score, HAQ-DI item score, Pain-VAS or DAS28 as explanatory variables (14, 15, 20, 60). Some researches would also include some demographic indicators as explanatory variables, like age and gender. However, we found that not only age and gender will affect the health utility values of RA patients, but also BMI will affect them. To enhance the accuracy of the mapping algorithm, we also tried to incorporate some clinical indicators, such as ESR, CRP, DAS28, PhGADA, PtAAP, and PtGADA. The correlations test showed that

EQ-5D-5L utility score had a strong negative correlation with the HAQ-DI score (−0.7067), a moderate negative correlation with PtAAP (−0.4040) and PtGADA (−0.4166), and a weak correlation with age, BMI, ESR, CRP, DAS28-ESR, DAS28-CRP, TJC, SJC, etc. Given the correlation between PtAAP and PtGADA was strong, and the correlation between PtGADA and EQ-5D-5L score was stronger than that between PtAAP and EQ-5D-5L score, we included the PtGADA as an explanatory variable. By the same token, we chose DAS28-ESR rather than DAS28-CRP. Due to the DAS28 was calculated by ESR, CRP, TJC, and SJC, we excluded the four indicators. Finally, HAQ-DI score, gender, age, BMI, PtGADA and DAS29-ESR were included as explanatory variables. Although the results showed that the more variables included, the more accurate the mapping algorithm was, we still recommended that researchers to choose a mapping algorithm based on actual data. And we provided mapping algorithms for different combinations of variables.

Among the seven statistics models we used, Beta model performed the best, followed by the MVROD-Probit. Some researches had presented that direct mapping with mixture model was better than direct mapping with linear regression, better than the response mapping (36, 52). ALDVMM was such a mixture model which could effectively capture the multimodal properties of EQ-5D utility score (18). In this research, however, ALDVMM had not performed as well as expected. The ALDVMM with 2 or more components suffered from the problem of non-convergence, which may change as the simple size increases. About MVROD-Probit, the results of Thomas' research showed that MVROD-Probit performed better than Beta, which was different from our results (13). Although our sample size was larger than that of Thomas' research, given the gap between the two sample size was small and our sample did not cover all the health status of patients, we thought a larger sample with more health status was necessary to demonstrate the difference of the two models.

Several limitations may affect the representativeness of our results. Firstly, despite the relatively rich source of the sample, the small sample size affected the using of model, like ALDVMM and MVROD-Probit, and increased the uncertainty of the results. Secondly, although the bootstrap was used to provide an assessment of the internal validity of the mapping procedures considered in this research, we could not verify the external validity of the mapping algorithm. Thirdly, most early RA patients have no typical clinical symptoms (61), and the treatment and economic evaluation mainly occur in these patients (62, 63), in China. Thus, our sample only included RA patients with middle and advanced stage, which may cause bias in predicting health utility values for patients with early stage. Finally, no patients chose 5 at mobility and anxiety/depressions dimensions of EQ-5D-5L in our sample, which may lead to bias for the results of response mapping.

Conclusion

More models and variables were used to construct the mapping algorithms in this research. And we found that mapping algorithms based on Beta performed better. Also, the more variables included, the mapping algorithm performed more better. Researchers could reasonably choose the combinations of variables and the mapping algorithm recommended based on the actual available data. These mapping algorithms could help researchers to obtain the health utility values of Chines RA patients and thus conduct other studies, like pharmacoeconomic evaluation.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Clinical Trial Ethics Committee of Huashan Hospital Affiliated to Fudan University (Reference Number 2019-252). The patients/participants provided their written informed consent to participate in this study.

Author contributions

CW, QW, and XX made their contributions to the conception and design of the work. CW, YH, and QW made their contributions to the acquisition and analysis of the data. ZX and YH made their contributions to the interpretation of data. CW made contributions to drafting of the work. CW, ZX, and XX made their contributions to revision of the work. All authors of this work has approved the submitted version and have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work.

Acknowledgments

The authors of this manuscript acknowledge that this article could not have been finished without the help of the many RA patients and their doctors involved in the course of data collection. Special tribute is also paid to the professionals who provided inspiring advice.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1123552/full#supplementary-material

References

1. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet.* (2016) 388:2023–38. doi: 10.1016/S0140-6736(16)30173-8

2. Association CR. Chinese guideline for the diagnosis and treatment of rheumatoid arthritis. *Chin J Internal Med.* (2018) 57:242–51. doi: 10.3760/cma.j.issn.0578-1426.2018.04.004

3. Tasiemski T, Angiaszwili-Biedna N, Wilski M. Assessment of objective and subjective quality of life in people with rheumatoid arthritis—Preliminary study. *Ortop Traumatol Rehabil.* (2009) 11:346–59.

4. Xinping T, Mengtao L, Xiaofeng Z. The challenges and opportunities for the management of rheumatoid arthritis in China: an annual report of 2019. *Chin J Internal Med.* (2021) 60:593–8. doi: 10.3760/cma.j.cn112138-20210207-00113

5. Xiao-feng Z, Song-lin Z, Ai-chun T, Xiao-ping X. Disease burden and quality of life of rheumatoid arthritis in china: a systematic review. *Chin J Evid Med.* (2013) 13:300–7.

6. Pan CW, Sun HP, Zhou HJ, Ma Q, Xu Y, Luo N, et al. Valuing health-related quality of life in type 2 diabetes patients in China. *Med Decis Making.* (2016) 36:234–41. doi: 10.1177/0272989X15606903

7. Liu G. *China Guidelines for Pharmacoeconomic Evaluations*. Beijing: China Market Press (2020).

8. Sun S, Stenberg E, Cao Y, Lindholm L, Salén KG, Franklin KA, et al. Mapping the obesity problems scale to the SF-6D: results based on the Scandinavian obesity surgery registry (SOReg). *Eur J Health Econ.* (2022) 3:1473. doi: 10.1007/s10198-022-01473-7

9. Spilker BA, editor. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia: Lippincott-Raven (1996).

10. Fayers PM, Machin D, editors. *Quality of Life: The Assessment, Analysis, and Interpretation of Patient-Reported Outcomes*. New Jersey: Wiley (2007).

11. Hasegawa H. Assessment of disease activity, structural damage, and function in rheumatoid arthritis. *Methods Mol Biol.* (2018) 1868:243–50. doi: 10.1007/978-1-4939-8802-0_25

12. Dexin Z, Yin L, Xijia T, Shanze W, Jianda M, Jianzi L, et al. Health utility measurement for rheumatoid arthritis: mapping HAQ-DI score onto the EQ-5D-5L utility score. *J Trop Med.* (2022) 22:295–300.

13. Patton T, Hu H, Luan L, Yang K, Li SC. Mapping between HAQ-DI and EQ-5D-5L in a Chinese patient population. *Qual Life Res.* (2018) 27:2815–22. doi: 10.1007/s11136-018-1925-1

14. Mlcoch T, Tuzil J, Sedova L, Stolfa J, Urbanova M, Suchy D, et al. Mapping quality of life (EQ-5D) from DAPsA, clinical DAPsA, and HAQ in psoriatic arthritis. *Patient.* (2018) 11:329–40. doi: 10.1007/s40271-017-0285-1

15. Kim HL, Kim D, Jang EJ, Lee MY, Song HJ, Park SY, et al. Mapping health assessment questionnaire disability index (HAQ-DI) score, pain visual analog scale (VAS), and disease activity score in 28 joints (DAS28) onto the EuroQol-5D (EQ-5D) utility score with the KORean Observational study Network for Arthritis (KORONA) registry data. *Rheumatol Int.* (2016) 36:505–13. doi: 10.1007/s00296-016-3427-1

16. Bujkiewicz S, Thompson JR, Sutton AJ, Cooper NJ, Harrison MJ, Symmons DP, et al. Use of Bayesian multivariate meta-analysis to estimate the HAQ for mapping onto the EQ-5D questionnaire in rheumatoid arthritis. *Value Health.* (2014) 17:109–15. doi: 10.1016/j.jval.2013.11.005

17. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. *Rheumatology.* (2013) 52:944–50. doi: 10.1093/rheumatology/kes400

18. Hernández Alava M, Wailoo AJ, Ara R. Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. *Value Health.* (2012) 15:550–61. doi: 10.1016/j.jval.2011.12.014

19. Carreño A, Fernández I, Badia X, Varela C, Roset M. Using HAQ-DI to estimate HUI-3 and EQ-5D utility values for patients with rheumatoid arthritis in Spain. *Value Health.* (2011) 14:192–200. doi: 10.1016/j.jval.2010.11.001

20. Versteegh MM, Rowen D, Brazier JE, Stolk EA. Mapping onto Eq-5 D for patients in poor health. *Health Qual Life Outcomes.* (2010) 8:141. doi: 10.1186/1477-7525-8-141

21. Ducournau P, Kielhorn A, Wintfeld NS. Comparison of linear and nonlinear utility mapping between HAQ and EQ-5D using pooled data from the tocilizumab trials option and lithe. *Rheumatology.* (2009) 48:I107–I8.

22. Tian F, Wen Z, Li J, Luo X, Deng L, Zhang L, et al. Cost-effectiveness of Anbainuo plus methotrexate compared to conventional disease-modifying antirheumatic drugs for rheumatoid arthritis patients in China. *Annals Translat Med.* (2021) 9:1165. doi: 10.21037/atm-21-3132

23. Tian F, Li JY, Wen ZH, Luo XW, Deng L, Zhang L, et al. A novel etanercept biosimilar Anbainuo plus methotrexate exhibits increased cost-effectiveness compared to conventional disease-modifying anti-rheumatic drugs in treating rheumatoid arthritis patients. *Medicine.* (2019) 98:e17750. doi: 10.1097/MD.0000000000017750

24. Shi ZC, Fei HP, Wang ZL. Cost-effectiveness analysis of etanercept plus methotrexate vs triple therapy in treating Chinese rheumatoid arthritis patients. *Medicine.* (2020) 99:e16635. doi: 10.1097/MD.0000000000016635

25. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European league against rheumatism collaborative initiative. *Arth Rheum.* (2010) 62:2569–81. doi: 10.1002/art.27584

26. Zou Y, Cheung PP, Teoh LK, Chen C, Lahiri M. Sociodemographic factors as determinants of disease, disability and quality of life trajectories in early rheumatoid arthritis: a multi-ethnic inception cohort study. *Int J Rheum Dis.* (2020) 23:55–64. doi: 10.1111/1756-185X.13747

27. Zhang L, Xia Y, Zhang Q, Fu T, Yin R, Guo G, et al. The correlations of socioeconomic status, disease activity, quality of life, and depression/anxiety in Chinese patients with rheumatoid arthritis. *Psychol Health Med.* (2017) 22:28–36. doi: 10.1080/13548506.2016.1198817

28. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheumatism.* (1995) 38:44–8. doi: 10.1002/art.1780380107

29. EQ-5D-5L | About: The EuroQol Group) (2021). Available online at: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/ (accessed May 9, 2022).

30. Luo N, Wang Y, How CH, Tay EG, Thumboo J, Herdman M. Interpretation and use of the 5-level EQ-5D response labels varied with survey language among Asians in Singapore. *J Clin Epidemiol.* (2015) 68:1195–204. doi: 10.1016/j.jclinepi.2015.04.011

31. Xia J, Wu NW, Ma TP Yu C, Li NX. Evaluation of reliability and validity of EQ-5D-5L based on residents in Southwest China. *Sichuan da xue xue bao Yi xue ban J Sichuan Univ Med Sci Ed.* (2020) 51:691–4.

32. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. *Value Health.* (2017) 20:662–9. doi: 10.1016/j.jval.2016.11.016

33. Bruce B, Fries JF. The stanford health assessment questionnaire: a review of its history, issues, progress, and documentation. *J Rheumatol.* (2003) 30:167–78.

34. The health assessment questionnaire (HAQ) disability index (DI) of the clinical health assessment questionnaire (version 96.4): National Institute of Environmental Health Sciences. Available online at: https://www.niehs.nih.gov/research/resources/assets/docs/haq_instructions_508.pdf (accessed May 9, 2022).

35. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR good practices for outcomes research task force report. *Value Health.* (2017) 20:18–27. doi: 10.1016/j.jval.2016.11.006

36. Yang F, Wong CKH, Luo N, Piercy J, Moon R, Jackson J. Mapping the kidney disease quality of life 36-item short form survey (KDQOL-36) to the EQ-5D-3L and the EQ-5D-5L in patients undergoing dialysis. *Eur J Health Econ.* (2019) 20:1195–206. doi: 10.1007/s10198-019-01088-5

37. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. The MAPS reporting statement for studies mapping onto generic preference-based outcome measures: explanation and elaboration. *Pharmacoeconomics.* (2015) 33:993–1011. doi: 10.1007/s40273-015-0312-9

38. Round J, Hawton A. Statistical alchemy: conceptual validity and mapping to generate health state utility values. *PharmacoEconomics - open.* (2017) 1:233–9. doi: 10.1007/s41669-017-0027-2

39. Correlation and regression: British Medical Journal. (2018). Available online at: https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression (accessed May 9, 2022).

40. Vilsbøll AW, Kragh N, Hahn-Pedersen J, Jensen CE. Mapping dermatology life quality index (DLQI) scores to EQ-5D utility scores using data of patients with atopic dermatitis from the national health and wellness study. *Qual Life Res.* (2020) 29:2529–39. doi: 10.1007/s11136-020-02499-1

41. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ.* (2010) 11:215–25. doi: 10.1007/s10198-009-0168-z

42. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess.* (2014) 18:1–224. doi: 10.3310/hta18090

43. Xu RH, Dong D, Luo N, Wong EL, Yang R, Liu J, et al. Mapping the Haem-A-QoL to the EQ-5D-5L in patients with hemophilia. *Qual Life Res.* (2022) 31:1533–44. doi: 10.1007/s11136-021-03051-5

44. Cuntong W. *Advanced Regression Analysis*. Beijing: Higher Education Press. (2017).

45. Chen G, Khan MA, Iezzi A, Ratcliffe J, Richardson J. Mapping between 6 multiattribute utility instruments. *Med Decis Making.* (2016) 36:160–75. doi: 10.1177/0272989X15578127

46. Yohai VJ. High breakdown-point and high efficiency robust estimates for regression. *Ann Stat.* (1987) 15:642–56. doi: 10.1214/aos/1176350366

47. Verardi V, Croux C. Robust regression in stata. *Stata J.* (2009) 9:439–53. doi: 10.1177/1536867X0900900306

48. Hawton A, Green C, Telford C, Zajicek J, Wright D. Using the Multiple Sclerosis Impact Scale to estimate health state utility values: mapping from the MSIS-29, version 2, to the EQ-5D and the SF-6D. *Value Health.* (2012) 15:1084–91. doi: 10.1016/j.jval.2012.07.007

49. Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making.* (2006) 26:401–9. doi: 10.1177/0272989X06290496

50. Basu A, Manca A. Regression estimators for generic health-related quality of life and quality-adjusted life years. *Med Decis Making.* (2012) 32:56–69. doi: 10.1177/0272989X11416988

51. Hunger M, Döring A, Holle R. Longitudinal beta regression models for analyzing health-related quality of life scores over time. *BMC Med Res Methodol.* (2012) 12:144. doi: 10.1186/1471-2288-12-144

52. Gray LA, Hernández Alava M, Wailoo AJ. Development of methods for the mapping of utilities using mixture models: mapping the AQLQ-S to the EQ-5D-5L and the HUI3 in patients with asthma. *Value Health.* (2018) 21:748–57. doi: 10.1016/j.jval.2017.09.017

53. Gray LA, Wailoo AJ, Hernandez Alava M. Mapping the FACT-B Instrument to EQ-5D-3L in patients with breast cancer using adjusted limited dependent variable mixture models vs. response mapping. *Value Health.* (2018) 21:1399–405. doi: 10.1016/j.jval.2018.06.006

54. Dixon P, Hollingworth W, Sparrow J. Mapping to quality of life and capability measures in cataract surgery patients: from Cat-PROM5 to EQ-5D-3L, EQ-5D-5L, and ICECAP-O using mixture modelling. *MDM policy & practice.* (2020) 5:2381468320915447. doi: 10.1177/2381468320915447

55. Xu RH, Wong ELY, Jin J, Dou Y, Dong D. Mapping of the EORTC QLQ-C30 to EQ-5D-5L index in patients with lymphomas. *Eur J Health Econ.* (2020) 21:1363–73. doi: 10.1007/s10198-020-01220-w

56. Wang K, Guo X, Yu S, Gao L, Wang Z, Zhu H, et al. Mapping of the acromegaly quality of life questionnaire to ED-5D-5L index score among patients with acromegaly. *Eur J Health Econ.* (2021) 22:1381–91. doi: 10.1007/s10198-021-01318-9

57. Conigliani C, Manca A, Tancredi A. Prediction of patient-reported outcome measures via multivariate ordered probit models. *J Royal Stat Soc Series* A. (2015) 178:12072. doi: 10.1111/rssa.12072

58. Zhang L, Chen F, Geng S, Wang X, Gu L, Lang Y, et al. Methotrexate (MTX) plus hydroxychloroquine versus MTX plus leflunomide in patients with MTX-resistant active rheumatoid arthritis: a 2-year cohort study in real world. *J Inflamm Res.* (2020) 13:1141–50. doi: 10.2147/JIR.S282249

59. Neilson AR, Jones GT, Macfarlane GJ, Pathan EM, McNamee P. Generating EQ-5D-5L health utility scores from BASDAI and BASFI: a mapping study in patients with axial spondyloarthritis using longitudinal UK registry data. *Eur J Health Econ.* (2022) 23:1357–69. doi: 10.1007/s10198-022-01429-x

60. Gorostiza I, Ansola L, Galindez E. Updated estimation of the EQ5D quality of life questionnaire utility values through HAQ-DI mapping for Spain. *Ann Rheum Dis.* (2017) 76:241–2. doi: 10.1136/annrheumdis-2017-eular.4836

61. Huajun X, Huimei Z. Ultrasonographic diagnosis of wrist joint disease in early stage of rheumatoid arthritis. *China Modern Doctor.* (2017) 55:106–10.

62. Ghabri S, Lam L, Bocquet F, Spath HM. Systematic literature review of economic evaluations of biological treatment sequences for patients with moderate to severe rheumatoid arthritis previously treated with disease-modifying anti-rheumatic drugs. *Pharmacoeconomics.* (2020) 38:459–71. doi: 10.1007/s40273-020-00887-6

63. Li J, Wen Z, Cai A, Tian F, Zhang L, Luo X, et al. Real-world cost-effectiveness of infliximab for moderate-to-severe rheumatoid arthritis in a medium-sized city of China. *J Comp Eff Res.* (2017) 6:205–18. doi: 10.2217/cer-2016-0086

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 04 May 2023
DOI 10.3389/fpubh.2023.1122718

OPEN ACCESS

EDITED BY
Ke Yan,
National University of Singapore, Singapore

REVIEWED BY
Marija Jevtic,
University of Novi Sad, Serbia
Wenxin Wang,
Shantou University, China

*CORRESPONDENCE
Huichao Han
✉ huichaoh@163.com
Chenxi Hai
✉ 20201244004@nuist.edu.cn

RECEIVED 13 December 2022
ACCEPTED 03 April 2023
PUBLISHED 04 May 2023

CITATION
Han H, Hai C, Wu T and Zhou N (2023) How does digital infrastructure affect residents' healthcare expenditures? Evidence from Chinese microdata.
Front. Public Health 11:1122718.
doi: 10.3389/fpubh.2023.1122718

How does digital infrastructure affect residents' healthcare expenditures? Evidence from Chinese microdata

Huichao Han*, Chenxi Hai*, Tianqi Wu and Nianchi Zhou

School of Business, Nanjing University of Information Science & Technology, Nanjing, China

Healthcare expenditure is only one of the heavy burdens that families face in developing countries. Current research mainly focuses on analyzing the effects of financial policy. There is a lack of studies that examine the understanding and assessment of the impact of digital infrastructure on this issue. In this study, we used the Broadband China policy as a quasi-natural experiment to explore the impact of digital infrastructure on residents' healthcare expenditures in China. Using the differences-in-differences (DID) model and micro-survey data, we found that digital infrastructure has a positive impact on reducing the burden of healthcare expenditure in China. Our findings indicate that residents in cities can save up to 18.8% on healthcare expenses following large-scale digital infrastructure construction. Through mechanism analysis, we found that digital infrastructure reduces residents' healthcare expenditures by improving both commercial insurance availability and the healthcare efficiency of residents. In addition, the effects of digital infrastructure on reducing healthcare expenditure are more pronounced among middle-aged individuals, those with low levels of education, and those with low incomes, which indicates this digital construction wave helps bridge the social gap between the poor and the rich. This study provides compelling evidence of the positive impact of digital society construction on social health and wellbeing.

KEYWORDS

healthcare expenditure, digital infrastructure, Broadband China, commercial insurance availability, residents' self-rated health

1. Introduction

Due to environmental pollution, an aging society, and the COVID-19 pandemic, healthcare expenditure is increasing in China and globally. According to World Health Organization (WHO) statistics, more than 500 million people have been pushed into extreme poverty due to the heavy burden of healthcare costs. This situation often leads to a decline in health status and decreased productivity among poorer groups (1, 2). The excessive burden of healthcare expenditures has a negative impact on human wellbeing, especially in developing countries. As the world's largest developing country, China faces various challenges, such as insufficient healthcare resources, uneven distribution of healthcare resources, severe air pollution, and an increasingly aging society, all of which contribute to the escalation of healthcare expenditure (3). In 2021, the average healthcare expenditure per capita was 5,348 RMB, which accounted for 8.8% of total personal expenditures. Although government healthcare spending continues to increase, it has not effectively reduced the individual healthcare burden, especially for rural residents and low-income groups (4).

In 2018, the average individual healthcare share of expenditure in China was 35.8%, which is significantly higher than the levels in the US (10.8%) and Japan (12.8%). Hence, reducing healthcare expenditures is crucial for promoting universal health coverage and achieving sustainable development goals (SDGs) in China.

Digital technology is widely considered a potential solution to alleviate the pressure on healthcare systems (5–7). In recent years, significant improvements in digital infrastructure have driven a deeper integration of digital networks within the realm of healthcare, thereby facilitating access to health and medical care information (8). Current studies have found that digital technology not only helps patients recover but also effectively mitigates the rise of chronic diseases (9). In addition, the widespread use of digital technology plays an important role in improving air quality (10, 11), which helps reduce medical and defensive expenditures for air pollution-induced diseases (12–15). Therefore, digital technology has the potential to alleviate the burden on healthcare expenditures, even though there is no direct causal relationship between the construction of digital infrastructure and healthcare expenditures. Therefore, assessing the social welfare effect of digital infrastructure is crucial.

Previous studies have researched the relevant factors of healthcare expenditure from different perspectives. The first type of study examines the effects of financial policies, such as medical insurance, medical assistance, and pensions, on reducing residents' healthcare expenditures (16, 17). Ma et al. found that access to a new rural social pension significantly reduced the proportion of individual medical expenses (18). The second category of study extensively discusses the impact of environmental pollution on healthcare expenditures. A number of studies have confirmed that environmental pollution increases healthcare expenditures (12, 19–21). Xia et al. (22) found that both higher air pollution levels and longer-duration pollution events significantly increased healthcare expenditure. Liao et al. (23) used microdata to quantify the effect of air pollution on healthcare expenditures. Third, socioeconomic factors such as industrial agglomeration and education level may also influence the healthcare expenditure of residents (24–26). Li et al. (27) found that a higher level of education significantly reduces the occurrence of catastrophic medical expenses.

In summary, the existing literature rarely explores healthcare expenditure from the perspective of technological change. Only a few studies have examined the impact of Internet applications on residents' healthcare expenditures (28). Moreover, most studies in the literature have failed to address endogeneity issues, such as omitted variables, which can create causal problems between residents' healthcare expenditure and the factors that influence it in current research.

This study attempted to bridge this gap by taking the Broadband China Pilot Policy as a quasi-natural experiment to assess the impact of digital infrastructure construction on residents' healthcare expenditures. The method has been proven to be an effective way of addressing endogeneity problems. We utilized unbalanced panel data by matching the 2010–2018 microdata of the China Family Panel Studies (CFPS) with Broadband China at the city level. Furthermore, we incorporated individual, household, and city-level characteristics associated with digital infrastructure to account for the effects of underlying factors on healthcare expenditure. This study provided two mechanisms: digital infrastructure reduces residents' healthcare expenditures by improving commercial insurance availability, and it also improves healthcare efficiency for residents, which, in turn, reduces the healthcare burden.

This study contributed to the current literature as follows. First, based on both individual-level microdata and city-level data, we examined the impact of digital infrastructure construction on healthcare expenditure from the perspective of technological change. Few studies have examined the direct healthcare effects of digital infrastructure construction, especially in the Chinese context. Second, to address the potential endogenous problem in the empirical study, we used the differences-in-differences (DID) method and employed multi-year unbalanced panel data with time and province fixed effects, controlling for individual, household, and city characteristics to reduce the bias of the estimates. Finally, as studies exploring the potential mechanisms of the impact of digital technology on healthcare expenditure are relatively limited, we identified two underlying mechanisms that explain the impact of digital infrastructure on healthcare expenditure.

The rest of the article is organized as follows. Section 2 provides a policy background. Section 3 introduces empirical model variables and data specifications. Section 4 presents the estimation results for the digital infrastructure on healthcare expenditure and a series of robustness tests, mechanism analyses, and heterogeneous effect studies among different demographic groups. Section 5 concludes the article.

2. Policy background

Since the 1990s, China has been promoting its broadband network coverage and enhancing information transmission speeds. However, despite significant progress, China still lags behind Western countries in terms of digital infrastructure. To accelerate China's digital construction, in August 2013, China launched the *Broadband China Strategy and Implementation Plan* (hereinafter referred to as "Broadband China" for abbreviation). The purpose of Broadband China was to select the pilot cities that would receive significant investments in digital infrastructure from the central and regional governments. The first batch of 39 pilot cities was named in 2014, with the second and third batches of cities later selected in 2015 and 2016, respectively. Since the implementation of the Broadband China strategy, China has made significant progress with respect to digital infrastructure. In 2020, fixed broadband access capacity in Chinese cities had generally exceeded 100 Mbps, and fiber-optic broadband had been made available to over 98% of villages. The proportion of fiber-access users in China had reached 93.2%, which is significantly higher than the OECD average level of 26.8%. Hence, it is fair to conclude that China is leading the world in digital infrastructure construction. Not only has China constructed the biggest 4G networks, but it is also expanding 5G networks. The implementation of the Broadband China strategy has significantly improved the level of China's digital infrastructure. Consequently, the Broadband China strategy provides a rare opportunity for quasi-natural

experimentation to assess the profound socioeconomic impact of digital infrastructure construction.

3. Methods

3.1. Methodology

In this study, we used the launch of the Broadband China strategy as a quasi-natural experiment to examine the impact of digital infrastructure on residents' healthcare expenditures. Considering that the policy of Broadband China was implemented in different years, we referred to Beck et al. (29) and constructed a time-varying DID model.

$$lnHE_{i,j,t} = \alpha_0 + \alpha_1 Policy_{i,j,t} + \sum \gamma_j X_{i,j,t} + \mu_i + \eta_j + \upsilon_t + \varepsilon_{i,j,t} \quad (1)$$

where ln $HE_{i,j,t}$ represents the healthcare expenditure of individual i in city j in year t; $Policy_{i,j,t}$ represents the Broadband China Pilot Policy; $X_{i,j,t}$ represents a set of control variables, μ_j represents city fixed effects, $\varepsilon_{i,j,t}$ represents the residual term, υ_t represents the fixed effect of the year, α_0 is the constant term, and α_1 and yj are the variable coefficients. This study evaluated the effect of the Broadband China Pilot Policy on healthcare expenditure by observing the significance and magnitude of the variable coefficient α_1.

3.2. Measure and description of variables

We mainly used two kinds of data. The key microdata on residents' healthcare expenditures in this study was collected from the China Family Panel Studies (CFPS), a nationally representative longitudinal survey of communities, households, and individuals launched in 2010 by the Institute of Social Science Survey (ISSS) at Peking University in China. CFPS is committed to providing the academic community with the most comprehensive and highest-quality survey data on contemporary China. The data used in this study were from CFPS 2010–2018. To obtain accurate estimation results, we collected two types of data from CFPS, including respondents' individual characteristics and household characteristics. This study considered variables that may be relevant to residents' healthcare expenditures, including information on respondents and households. The CFPS consists of three main components: an adult database, a household database, and a community database. However, they are separate from each other. If we want to control for both individual and household-level variables, we need to merge them through unique household codes. In addition, the policy shocks we used are at the prefecture and city levels; thus, we also needed to merge the already-merged CFPS data with the city-level data through unique city codes.

We also used macrodata, which mainly included GDP per capita, expenditure on science and education, population density, urban green coverage, and the value-added of secondary industries for each prefecture-level city in China from 2010 to 2018. All macrodata are taken from the China Urban Statistical Yearbook.

The sample selection process for this study was as follows: First, we matched individual and household data in the CFPS by unique household codes to obtain the CFPS dataset. Then, we matched the CFPS dataset with the municipality data using the unique municipality code. With the previous processing, we obtained a dataset covering individual, household, and prefecture-level characteristics.

3.2.1. Dependent variables

This study considered the Broadband China Pilot Policy to be a quasi-natural experiment and used it to measure digital infrastructure construction. The Ministry of Industry and Information Technology and the National Development and Reform Commission of China designated 119 Broadband China demonstration cities in 2014, 2015, and 2016. We adopted the form of policy using a dummy variable; the variable *Policy* equals 1 if the city j was selected as the pilot city from the year t. Otherwise, it equaled 0.

3.2.2. Independent variables

We used CFPS survey data to determine the residents' healthcare expenditures. In the CFPS questionnaires, a special question was asked: "How much has your household spent on healthcare in the past year?" The expenditure was measured by the constant price. We considered the natural logarithm of the variable's value.

3.2.3. Mechanism variables

We examined two mechanism variables. The first was the ease of purchasing commercial insurance. It required interviewees to answer, "How much does your family spend on commercial insurance?" The mechanism variable was coded as 1 if the interviewer had bought commercial insurance and 0 otherwise. The second mechanism variable was residents' health status, which was obtained from the CFPS questionnaire item: "Do you consider yourself to be in good health?" The answers ranged from 1 (worst health) to 5 (best health). To facilitate analysis, we recorded the response options from 1 to 3, with 1 indicating poor health, 3 indicating very good health, and 2 indicating average health.

3.2.4. Controlling variables

Two types of variables were controlled for in the analysis. The first type comprised individual demographic characteristics, which mainly include age, Hukou (household registration), gender, years of education, marital status, smoking and drinking status, household income per capita, household water source, and family size. The second comprised prefecture-level characteristics., which include population density (PD), science and education expenses (RD), greenery rate (Green), industrial structure (Second), and GDP per capita (GDPP). We applied the natural logarithm of the variables of household income per capita, PD, RD, Green, Second, and GDPP.

10.3389/fpubh.2023.1122718

TABLE 1 Descriptive statistics.

Variables	Observations	Mean	S.D.	Min	Max
HE	173,119	6.863	2.625	0	14.00
Policy	118,776	0.191	0.393	0	1
Age	173,893	45.88	17.27	16	110
Hukou	164,895	0.272	0.445	0	1
Gender	172,268	0.495	0.500	0	1
Education	168,057	7.277	4.826	0	23
Marriage	163,973	0.726	0.446	0	1
Smoke	161,652	0.284	0.451	0	1
Drink	135,298	0.144	0.351	0	1
Household water	176,539	0.660	0.474	0	1
Preincome	165,889	9.188	1.380	0.182	15.22
Family size	174,833	4.332	1.968	1	26
Status	158,392	2.916	1.018	1	5
PD	118,776	6.378	1.243	1.630	11.06
RD	118,776	10.35	1.390	7.116	15.53
Green	118,776	3.655	0.405	0.47	6.121
GDPP	118,776	10.56	0.571	8.773	12.16
Second	118,776	3.835	0.223	2.757	4.41
Insurance	176,539	0.405	0.491	0	1
Healthy	176,068	2.552	0.746	1	3

3.3. Data sources and descriptive statistics

Broadband information about China's pilot cities is issued by the Ministry of Industry and Information Technology and the National Development and Reform Commission. Healthcare expenditure mechanism Variables and individual-level control data come from the China Family Panel Studies (CFPS). CFPS is a national survey program initiated in 2010 that collects data from 25 provinces in China, covering 95% of the Chinese population. The sampling method for CFPS is based on a multi-stage approach. The CFPS program collects data every 2 years and aims to investigate family and individual information on a range of topics, including economic status, educational background, work status, and physical and mental health. The remaining controls in prefecture-level cities are from the China City Statistical Yearbook.

We matched healthcare expenditure and individual demographic variables with Broadband China and prefecture-level variables for each year to obtain a valid unbalanced panel data sample from 2010 to 2018. Table 1 shows the descriptions of the data. It includes observations, the mean, the standard deviation, and the maximum and minimum values of the main variables.

TABLE 2 Estimation results of the benchmark model.

Variables	HE		
	(1)	(2)	(3)
Policy	−0.121*** (0.030)	−0.147*** (0.037)	−0.188*** (0.039)
Age		0.003 (0.030)	0.002 (0.029)
Hukou		0.110 (0.077)	0.108 (0.077)
Gender		−0.456 (0.382)	−0.450 (0.381)
Education		−0.019 (0.012)	−0.019 (0.012)
Marriage		−0.045 (0.063)	−0.043 (0.063)
Smoke		−0.197*** (0.052)	−0.200*** (0.052)
Drink		−0.096** (0.043)	−0.096** (0.043)
Lifewater		−0.131*** (0.034)	−0.139*** (0.034)
Preincome		0.089*** (0.012)	0.087*** (0.012)
Familysize		0.179*** (0.011)	0.180*** (0.011)
Status		−0.009 (0.013)	−0.009 (0.013)
PD			0.002 (0.022)
RD			0.047** (0.023)
Green			0.140*** (0.030)
GDPP			0.385*** (0.116)
Second			−0.026 (0.176)
Constant	5.878*** (0.436)	4.101** (1.891)	−1.091 (2.113)
Individual FEs	Yes	Yes	Yes
Province FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	116,364	78,601	78,601
R-squared	0.012	0.022	0.023

In parentheses are standard errors adjusted for heteroskedasticity and clustering by both individual and year. *, **, and *** are significant at the levels of 10, 5, 1%, respectively.

4. Results

4.1. Baseline regression results

The baseline results of the impact of digital infrastructure on healthcare expenditures are reported in Table 2. Two-way fixed effects were controlled for in the main analysis. Column 1 reports the estimations of the impact of digital infrastructure on healthcare expenditures without controls, and Columns 2–3 display the regression results of models with individual characteristic controls

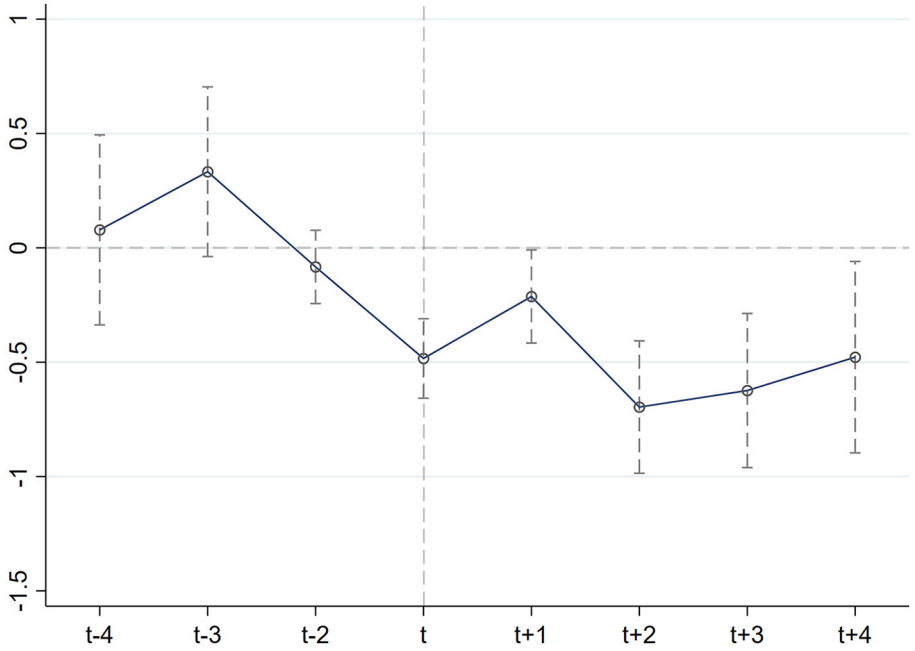


FIGURE 1
Plot of parallel trend test results.

and characteristic city controls introduced step by step. The results show that the coefficients of digital infrastructure are negative and significant at the 1% level. In other words, the digital infrastructure significantly reduces healthcare expenditure, whether controlled variables are added or not. The healthcare expenditure of treatment groups is, on average, reduced by 18.8% more than that of control groups. Therefore, digital infrastructure has a significant impact on reducing residents' healthcare expenditures.

4.2. Robustness tests

4.2.1. Parallel trend tests

Adopting the difference-in-differences model hinges on passing a parallel trend test; that is, the trend in healthcare spending by residents in pilot and non-pilot cities of Broadband China is the same as before the policy was taken. Following the method of Lyu et al. (30), we used an event-study approach to estimate the dynamic treatment effects in Broadband China. The empirical model is as follows:

$$lnHE_{i,j,t} = \theta_0 + \sum_{\tau=-1}^{\tau=-6} \theta_\tau pre_{i,j,t} + \theta_1 dummy_{i,j,t} + \sum_{\sigma=1}^{\sigma=5} \theta_\sigma post_{i,j,t} + \theta_2 control_{i,j,t} + \mu_i + \eta_j + \gamma_t + \varepsilon_{i,j,t} \quad (2)$$

where $lnME_{i,j,t}$ represents residents' healthcare expenditure, and *pre* is a set of counterfactual dummy variables. If it is assumed that the pilot policy of Broadband China has changed from τ implemented in (τ = 2,012, 2,010), then *pre* = 1, for the other years *pre* = 0. Assuming that the pilot policy of Broadband China was implemented since the σ in the year of implementation, *post* = 1, in other years *post* = 0, *dummy* = 1 in the year of implementation of the Broadband China policy, otherwise *dummy* = 0 in other years.

Figure 1 presents the results of the parallel trend test. The estimation results prior to implementing Broadband China were not significant. This result shows that prior to the introduction of Broadband China, there was no systematic difference between the treatment group and the control group. Since the beginning of Broadband China, the residents' healthcare expenditure in the treatment group has been significantly reduced. The sample satisfies the parallel trend assumption.

4.2.2. PSM-DID method

To overcome the systematic differences in the trends between the pilot cities of Broadband China and other cities and to reduce the estimation bias of the double difference method, this paper further uses the PSM-DID method to conduct robustness checks. Specifically, this study drew on the study by Heyman et al. to use the control variables in the benchmark regression as covariates (31). The samples were matched year by year using a matching method and then merged vertically with the matched data for each year to create a dataset that generated panel data for regression. The balance of the matching data (Figure 2) was checked. As shown in Figure 2, the deviation of the standardized mean of all matched variables after matching was <20%. This indicates that there was no systematic difference between the treatment group and the control group before the policy impact. The coefficient of PSM-DID in Table 3 (1)

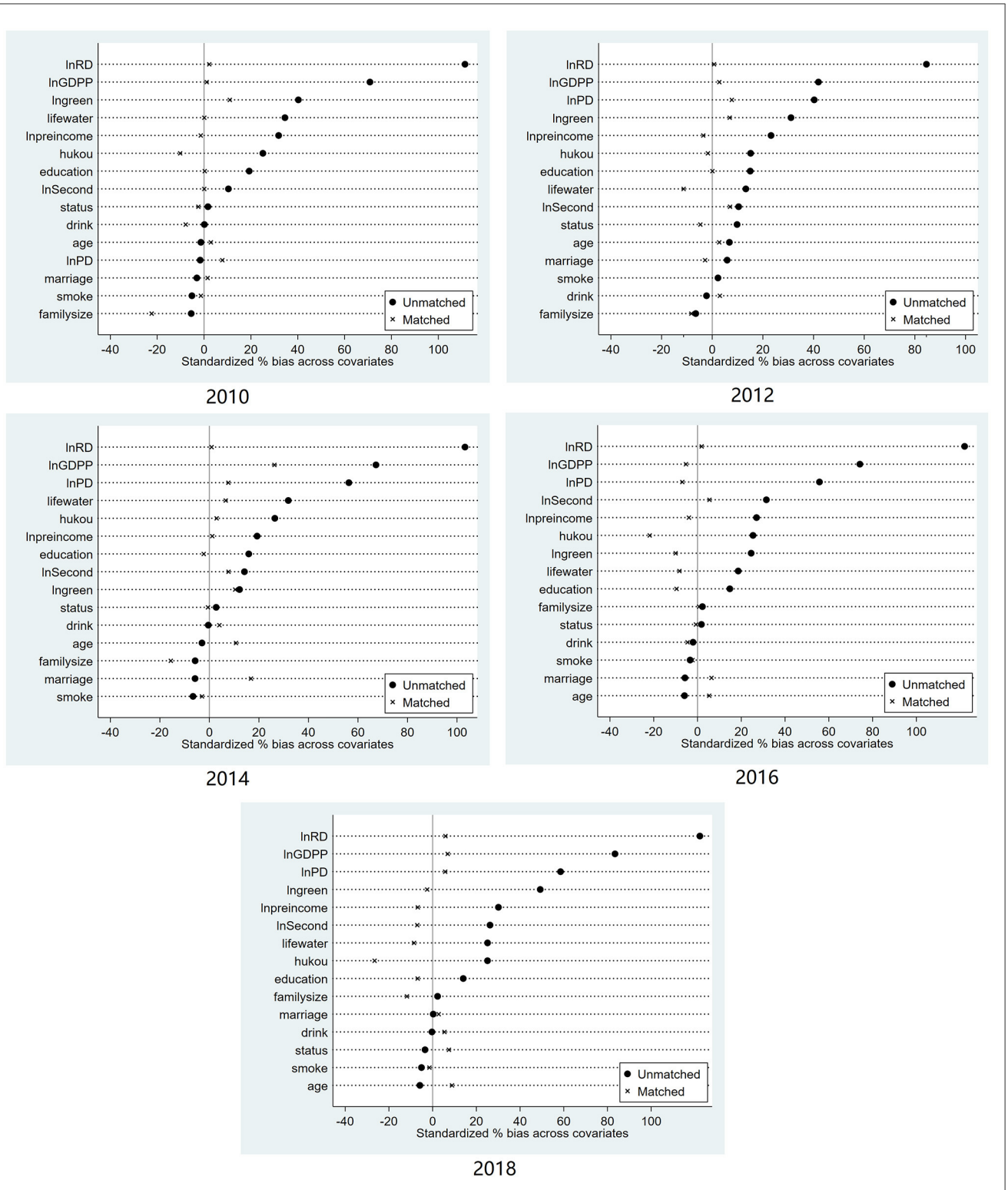


FIGURE 2
Balance test.

is −0.111 and significant. There is no significant difference when compared with the benchmark regression results, which further supports the empirical conclusion that the implementation of digital infrastructure has significantly reduced residents' healthcare expenditures.

4.2.3. Controlling other policies

Some other policies may exist that affect healthcare expenditure when Broadband China is implemented. The Broadband China dummy variable in the baseline regression model may include other policy shocks, which may lead to bias in the estimation results.

TABLE 3 Robustness test results.

Variables	ME				
	(1)	(2)	(3)	(4)	(5)
Policy	−0.111** (0.047)	−0.193*** (0.039)	−0.187*** (0.040)		−0.183*** (0.039)
Bigdata		−0.104*** (0.039)			
Smartcity			−0.008 (0.039)		
Fakepolicy				0.016 (0.388)	
Constant	1.163 (2.285)	−0.666 (2.115)	−1.096 (2.113)	2.392 (2.858)	2.359 (1.993)
Controls	Yes	Yes	Yes	Yes	Yes
Individual FEs	Yes	Yes	Yes	Yes	Yes
Province FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Observations	56,053	78,601	78,601	32,799	77,734
R-squared	0.023	0.023	0.023	0.022	0.023

In parentheses are standard errors adjusted for heteroskedasticity and clustering by both individual and year. *, **, and *** are significant at the levels of 10, 5, 1%, respectively.

We checked and selected some other policies to test whether they affect the effect of Broadband China. First, the Smart City Policy was proposed in 2012. Through a series of measures, this policy aimed to provide both a better living standard and better working services for citizens, create a more favorable business development environment for enterprises, and optimize the government with more efficient operation and management mechanisms. Among them, smart healthcare construction may also be beneficial in reducing healthcare expenditures for residents. Second, in 2016, China's National Development and Reform Commission, the Ministry of Industry and Information Technology, and the Central Internet Information Office issued a letter approving the establishment of a national-level comprehensive big data pilot zone. The Big Data Pilot Zone carries out systematic experiments around seven major tasks, including data resource management and sharing and opening, data center integration, data resource application, data element circulation, big data industry clustering, and big data system innovation. The policy may have an impact on residents' healthcare expenditures. Based on these considerations and to mitigate the potential impact of the Smart Cities Policy and Big Data pilot zone on the estimated results, we set the dummy variables for the Smart Cities in 2012, 2013, and 2014 and the Big Data pilot zone in 2016, respectively. We then introduced them into the baseline model together with Broadband China. Columns 1–3 of Table 3 show the results. The estimates show that Broadband China policy still significantly reduces healthcare expenditures after controlling for potential policy disruptions of smart cities and big data plot zones.

4.2.4. Counterfactual tests

The use of the DID model requires that the treatment and control groups be comparable. Without the implementation of the digital infrastructure, there would have been no significant difference in healthcare expenditure between the treatment and control groups due to changes over time. However, in addition to broadband in China, some other policies or random factors may also cause differences in healthcare expenditures. Although these differences are not associated with the construction of broadband in China, they may ultimately contribute to the conclusions drawn in the previous section. To rule out this possibility, we applied a strategy of changing the time by drawing on the method of Rao et al. (32). We used a sham experiment with a hypothetical policy shock in 2012 to examine whether those healthcare expenditures also differed between the treatment and control groups before and after 2012. Finding a significant negative effect empirically meant that the previous regression was not meaningful. Column 4 of Table 3 shows the estimated coefficient is not significant and also suggests that the results of the baseline regressions are not due to regular random factors.

4.2.5. Excluding first-tier cities

Due to the high level of digital infrastructure construction in the four first-tier cities of Beijing, Shanghai, Guangzhou, and Shenzhen, not only are their urban patterns and economic levels different from those of other cities, but there are also differences in economic decision-making and urban planning. Drawing on the method by Wu et al. (33), we excluded the samples from the four first-tier cities. Column 5 of Table 3 shows that the estimated coefficient is also significant and that digital infrastructure can still reduce residents' healthcare expenditures.

4.3. Heterogeneity analysis

4.3.1. Residents' age

Age has long been regarded as one of the critical factors in healthcare expenditure (34). We divided ages into three categories according to the World Health Organization (WHO) criteria, namely the young group (16–44 years old), the middle-aged group (45–60 years old), and the older adult group (over 60 years old). Excluding the age variable from the baseline regression model, the regression was estimated separately for each group. The results

10.3389/fpubh.2023.1122718

TABLE 4 Results of heterogeneity analysis.

Panel A: age			
Variables	HE		
	(1) 16–44	(2) 45–60	(3) >60
Policy	−0.075 (0.065)	−0.382*** (0.076)	−0.077 (0.084)
Constant	−5.419*** (2.031)	−4.230** (1.819)	2.203 (1.961)
Controls	Yes	Yes	Yes
Individual FEs	Yes	Yes	Yes
Province FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	34,986	25,869	17,747
R-squared	0.023	0.03	0.026

Panel B: education				
Variables	HE			
	(1) <6	(2) 6–9	(3) 9–12	(4) >12
Policy	0.207*** (0.062)	−0.245*** (0.069)	−0.212** (0.101)	0.118 (0.129)
Constant	−8.911*** (2.339)	10.830*** (4.199)	8.789** (3.482)	1.572 (3.899)
Controls	Yes	Yes	Yes	Yes
Individual FEs	Yes	Yes	Yes	Yes
Province FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	30,979	26,339	13,211	10,350
R-squared	0.026	0.022	0.033	0.036

Panel C: income level and region				
Variables	HE			
	(1) Low-income	(2) High-income	(3) Urban	(4) Rural
Policy	0.156** (0.071)	−0.105* (0.058)	−0.272*** (0.046)	0.026 (0.076)
Constant	0.448 (2.160)	3.104 (3.352)	0.289 (2.296)	2.065 (4.180)
Control	Yes	Yes	Yes	Yes
Individual FEs	Yes	Yes	Yes	Yes
Province FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	39,183	45,067	57,125	21,476
R-squared	0.023	0.026	0.024	0.024

In parentheses are standard errors adjusted for heteroskedasticity and clustering by both individual and year. *, **, and *** are significant at the levels of 10, 5, 1%, respectively.

of the estimations in Panel A of Table 4 indicate that digital infrastructure has no significant effect on the medical expenditure of residents under 44 years old and over 60 years old, while it has a significant negative effect on residents aged 45–60 years old. The reason for the result is that residents under 44 years generally have better health and lower healthcare expenditures, while people over 60 years have limited ability to benefit from digital infrastructure due to technological barriers. In contrast, for residents aged 45–60 years, medical expenditure tends to increase with age, making the digital infrastructure a significant negative factor impacting their healthcare expenditure level.

4.3.2. Residents' educational levels

Educational level is an important factor that affects healthcare expenditures. To account for potential heterogeneity in healthcare expenditure among residents with different educational levels, this study divided the residents into four groups based on their education level: low education level (6 years and below), a medium-low education level (6–9 years), a medium-high education level (9–12 years), and high education level (12 years and above). Panel B of Table 4 shows the heterogeneous estimated results of healthcare expenditure. The effect of digital infrastructure is significant for residents with an education level below 12 years. The above results

indicate that digital infrastructure provides more online platforms and information channels, making it easier for residents with low, medium-low, and medium-high education levels to acquire health knowledge to protect themselves against diseases and thus reduce medical expenses. Therefore, the digital infrastructure plays a significantly larger role in groups with low, medium-low, and medium-high education levels.

4.3.3. Residents' income levels

We analyzed the heterogeneity of income levels. Columns 1–2 in Panel C of Table 4 show that digital infrastructure has a significant negative effect on the healthcare expenditures of both low- and high-income residents. However, there is a noticeable difference in effectiveness between the two income groups, even though both remain significant at the conventional level. There are two possible reasons for the above heterogeneous results by income. First, residents with high incomes may pay more attention to healthcare and take more preventive actions to avoid risk than those with low incomes. Therefore, they are less affected by the digital infrastructure. Second, compared with the high-income group, digital infrastructure may increase low-income individuals' awareness of healthcare protection, leading them to be more affected by the construction of digital infrastructure.

4.3.4. Residents' Hukou

Considering the development differences across urban and rural areas, we further examined whether the effect of healthcare expenditure varies across different types of Hukou. The samples were divided into two groups: urban and rural. Columns 3–4 in Panel C of Table 4 present the estimated results for Hukou, showing that rural Hukou healthcare expenditure is not significantly negatively affected by digital infrastructure. However, in the case of urban Hukou, digital infrastructure may lead to lower healthcare expenditures for people. In summary, the above results also suggest that in urban regions with higher levels of digital infrastructure, digital infrastructure greatly benefits people's healthcare expenditures.

Overall, the heterogeneous effects of broadband vary by age, education, income, and region. These effects are mainly observed in middle-aged urban residents and people with low income and educational levels.

4.4. Mechanism analysis

What mechanisms explain the digital infrastructure and the reduction in residents' healthcare expenditures? In this section, we explored two channels through which digital infrastructure reduces the healthcare expenditures of residents: the accessibility of purchasing commercial insurance and the residents' health.

4.4.1. Accessibility of purchasing commercial insurance

Not only has digital infrastructure accelerated the application of new-generation digital technology, such as big data, artificial intelligence, and cloud computing, in many fields, but Internet insurance has also gradually emerged and become popular. Compared with traditional insurance, Internet insurance has the characteristics of convenience, timeliness, efficiency, innovation, and a small amount of high frequency, all of which can effectively reduce both transaction costs and the asymmetry of insurance information, breaking the spatial distance limitations and increasing the accessibility of commercial insurance for residents (35). Moreover, premium income from personal insurance (such as health, life, and accident insurance) accounts for over 80% of China's commercial insurance premium income. This indicates that commercial insurance is positively correlated with health insurance, which will reduce residents' healthcare expenditures. Therefore, digital infrastructure can reduce residents' healthcare expenditures by increasing the availability of commercial insurance.

4.4.2. Residents' health

A number of relevant studies confirm that Internet users have better physical and mental health (36–40) and that health is closely related to healthcare costs (41). Hence, we argued that digital infrastructure may reduce healthcare expenditure by improving residents' health. First, digital infrastructure promotes the utilization of the Internet in the medical field, improving residents' health by alleviating the uneven allocation of medical resources and improving treatment efficiency and medical services (42, 43). For instance, the application of real-time and virtual dialogue digital technology breaks the time and space constraints of medical services and reduces irrational medical treatment behavior, which not only improves residents' own health but also reduces medical expenses. Second, digital infrastructure is conducive to popularizing health knowledge (44), which helps residents choose healthier lifestyles, improve their own health (45), and ultimately reduce medical expenses.

Based on the analysis of the theoretical mechanism, this study considered the accessibility of purchasing commercial insurance and residents' health to be the mechanism variables for the digital infrastructure and residents' healthcare expenditure. Drawing on Chen et al. (45), we constructed the following model to validate the study mechanism.

$$M_{i,j,t} = \beta_0 + \beta_1 Policy_{i,j,t} + \sum \gamma_j X_{i,j,t} + \mu_j + \eta_j + \upsilon_t + \varepsilon_{i,j,t}, \quad (3)$$

where the mediating variable $M_{i,j,t}$ is a potential mechanism variable. The signs and significance of β_1 is the focus of this study.

Column 1 of Table 5 shows that the regression coefficient of digital infrastructure strategy on commercial insurance purchase is 0.039, which is significant at the 1% level. This suggests that network infrastructure construction increases residents' accessibility to purchasing commercial insurance, reducing their healthcare expenditure. The results in column 2 of Table 5 show that the regression coefficient of the digital infrastructure is 0.016, which is significant at the 10%

TABLE 5 Mechanism analysis.

Variables	Insurance	Health
	(1)	(2)
Policy	0.039*** (0.006)	0.016* (0.009)
Constant	0.462* (0.251)	2.083*** (0.445)
Controls	Yes	Yes
Individual FEs	Yes	Yes
Province FEs	Yes	Yes
Year FEs	Yes	Yes
Observations	79,111	79,103
R-squared	0.165	0.060

In parentheses are standard errors adjusted for heteroskedasticity and clustering by both individual and year. *, **, and *** are significant at the levels of 10, 5, 1%, respectively.

level, indicating that the network infrastructure significantly improves the health level of the residents and then reduces healthcare expenditures.

4.5. Social benefit analysis

The above empirical study demonstrates that digital infrastructure development significantly reduces residents' healthcare expenditures by increasing the accessibility of commercial insurance and enhancing the health of the population. In this section, we take a step further to estimate the social benefits.

Following Liao et al. (23) and Chen et al. (45), a cost-benefit analysis was conducted in this section to explore the total social costs and welfare benefits caused by digital infrastructure. The estimates in Table 2 suggest that the implementation of digital infrastructure reduces healthcare expenditures by 18.8 percent. Thus, the total social welfare benefit led by the digital infrastructure can be calculated by multiplying the estimated effect of digital infrastructure on healthcare expenditure by the total population size in China for each year and the annual average personal healthcare expenditure. For example, the product of the average healthcare expenditure (1,307 RMB), the total population size in 2016 (1.38 billion), and the estimated effect of digital infrastructure on healthcare expenditure (18.8%) is ∼325 million RMB. This means that digital infrastructure construction has reduced healthcare expenditure for society by 325.3 million RMB (or 46.9 million USD). Using a similar method, we can calculate that the construction of digital infrastructure reduced social healthcare expenditure by ∼419 million RMB (or 63.4 million USD) in 2018. We observed that the social benefits are becoming larger over time.

There should be other indirect social benefits except for the above direct social benefits. Since digital infrastructure plays a basic role in a digital society, the construction of digital infrastructure also has sweeping impacts on many aspects of society. For example, through online education, remote families living in some mountainous areas can get more quality education, improving human resources and bringing economic results such as a higher household income. Given the lack of relevant data, it is challenging to calculate such indirect benefits in healthcare. However, this area deserves research in the future.

5. Conclusion and implications

Given the context of rapid digital infrastructure construction and rising healthcare expenditure, it is important to leverage the potential of digital infrastructure to reduce healthcare expenditure and further improve residents' quality of life. Taking advantage of the quasi-natural experiment provided by the Broadband China policy and using a sample of CFPS from 2010 to 2018, our DID models show that digital infrastructure leads to a decrease in residents' healthcare expenditure.

The main conclusions are as follows. First, we found that digital infrastructure construction significantly reduces residents' healthcare expenditures. Compared with non-Broadband China pilot cities, the residents in Broadband China pilot cities reduced their healthcare expenditures by 18.8%, illustrating the apparent impact of sweeping digital technology advances on the residents' healthcare behavior. Second, the heterogeneous results show that the effects of digital infrastructure on reducing healthcare expenditure are higher for middle-aged residents, those with lower levels of education and low income, as well as those living in urban areas. Third, healthcare expenditure is influenced by digital infrastructure through the two underlying mechanisms of commercial insurance accessibility of purchasing and residents' health. Finally, we calculated the social welfare brought about by digital infrastructure. We estimated that the construction of digital infrastructure could reduce social healthcare expenditures by ∼419 million RMB, or 63.4 million USD, at 2018 exchange rates.

Although our research was based on Chinese data, it has worldwide implications. In contemporary times, finding new solutions to address social healthcare issues is urgent. Governments in both developed and developing countries often face financial constraints, especially in the current era of the pandemic, inflation, and other uncertainties. This research provided new ideas for overcoming these challenges.

Several policy implications can be derived from the results of this study. First, the government should accelerate the construction of digital infrastructure and promote the widespread application of digital technologies such as 5G in the healthcare sector. Digital infrastructure construction is a key influencing factor in the development of smart healthcare, which provides diverse access to medical treatment, improves treatment efficiency, reduces unreasonable medical practices, and ultimately reduces medical expenditures. This provides a solution for developing countries to achieve good health and wellbeing through sustainable development goals. Second, the government should be aware of the potential digital divide resulting from the construction of digital infrastructure. Efforts should be made to narrow this divide by improving the digital literacy of less educated individuals and older adults and by increasing investment in the construction of digital infrastructure in less developed areas such as rural areas. This

has significant implications for the digitalization of developing countries. Finally, the government should encourage insurance companies to use both the Internet and digital technology to provide diversified health insurance products, simplify the insurance purchase and claims process, effectively perform the insurance protection function, and ultimately lead to a reduction in healthcare expenses.

Despite these strengths, our study also has some limitations. First, due to data restrictions, we used the "Broadband China" policy as a proxy variable for digital infrastructure development, instead of directly measuring digital infrastructure development in cities. Second, this study measures the level of healthcare expenditure using total healthcare expenditure without differentiating between out-of-pocket and reimbursement costs. Future studies could use more detailed data to investigate the impact of digital infrastructure development on healthcare expenditures.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

Conceptualization, methodology, software, validation, resources, data curation, writing—original draft preparation, and literature collection: CH. Writing—review and editing: HH. Supervision: HH, TW, and NZ. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Shahnazi R, Dehghan Shabani Z. The effects of spatial spillover information and communications technology on carbon dioxide emissions in Iran. *Environ Sci Poll Res*. (2019) 26:24198–212. doi: 10.1007/s11356-019-05636-7

2. Onoka CA, Onwujekwe OE, Hanson K, Uzochukwu BS. Examining catastrophic health expenditures at variable thresholds using household consumption expenditure diaries. *Trop Med Int Health*. (2011) 16:1334–41. doi: 10.1111/j.1365-3156.2011.02836.x

3. Bai P, Tang Y, Zhang W, Zeng M. Does economic policy uncertainty matter for healthcare expenditure in China? a spatial econometric analysis. *Front Public Health*. (2021) 9:673778. doi: 10.3389/fpubh.2021.673778

4. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 years of healthcare reform in China: progress and gaps in universal health coverage. *Lancet*. (2019) 394:1192–204. doi: 10.1016/S0140-6736(19)32136-1

5. Gope P, Gheraibia Y, Kabir S, Sikdar B. A secure IoT-based modern healthcare system with fault-tolerant decision making process. *IEEE J Biomed Health Inform*. (2020) 25:862–73. doi: 10.1109/JBHI.2020.3007488

6. Zhu N, Diethe T, Camplani M, Tao L, Burrows A, Twomey N, et al. Bridging e-health and the internet of things: the sphere project. *IEEE Intell Syst*. (2015) 30:39–46. doi: 10.1109/MIS.2015.57

7. Chang S-H, Chiang R-D, Wu S-J, Chang W-T. A context-aware, interactive M-health system for diabetics. *IT Profess*. (2016) 18:14–22. doi: 10.1109/MITP.2016.48

8. Minetaki K, Akematsu Y, Tsuji M. Effect of e-health on medical expenditures of outpatients with lifestyle-related diseases. *Telemed e-Health*. (2011) 17:591–5. doi: 10.1089/tmj.2011.0019

9. Fan YJ, Yin YH, Da Xu L, Zeng Y, Wu F. IoT-based smart rehabilitation system. *IEEE Trans Indus Inform*. (2014) 10:1568–77. doi: 10.1109/TII.2014.2302583

10. Ozcan B, Apergis N. The impact of internet use on air pollution: evidence from emerging countries. *Environ Sci Pollut Res*. (2018) 25:4174–89. doi: 10.1007/s11356-017-0825-1

11. Magazzino C, Porrini D, Fusco G, Schneider N. Investigating the link among ICT, electricity consumption, air pollution, and economic growth in EU countries. *Energy Sources Part B Econ Plann Policy*. (2021) 16:976–98. doi: 10.1080/15567249.2020.1868622

12. Zhang J, Mu Q. Air pollution and defensive expenditures: evidence from particulate-filtering facemasks. *J Environ Econ Manag*. (2018) 92:517–36. doi: 10.1016/j.jeem.2017.07.006

13. Zhang H, Niu Y, Yao Y, Chen R, Zhou X, Kan H. The impact of ambient air pollution on daily hospital visits for various respiratory diseases and the relevant medical expenditures in Shanghai, China. *Int J Environ Res Public Health*. (2018) 15:425. doi: 10.3390/ijerph15030425

14. Ito K, Zhang S. Willingness to pay for clean air: evidence from air purifier markets in China. *J Polit Econ*. (2020) 128:1627–72. doi: 10.1086/705554

15. Chang TY, Huang W, Wang Y. Something in the air: pollution and the demand for health insurance. *Rev Econ Stud*. (2018) 85:1609–34. doi: 10.1093/restud/rdy016

16. Okoroh J, Essoun S, Seddoh A, Harris H, Weissman JS, Dsane-Selby L, et al. Evaluating the impact of the national health insurance scheme of Ghana on out of pocket expenditures: a systematic review. *BMC Health Serv Res*. (2018) 18:426. doi: 10.1186/s12913-018-3249-9

17. Kranker K. Effects of Medicaid disease management programs on medical expenditures: evidence from a natural experiment in Georgia. *J Health Econ*. (2016) 46:52–69. doi: 10.1016/j.jhealeco.2016.01.008

18. Ma C, Li S, Sun T, Tang R. The role of pensions in relieving rural residents' health care burden - why subsidizing income is better than subsidizing health insurance. *China Indus Econ*. (2021) 4:43–61.

19. Xue T, Zhu T, Peng W, Guan T, Zhang S, Zheng Y, et al. Clean air actions in China, PM2. 5 exposure, and household medical expenditures: a quasi-experimental study. *PLoS Med*. (2021) 18:e1003480. doi: 10.1371/journal.pmed.1003480

20. Monzón A, Guerrero Ma-J. Valuation of social and health effects of transport-related air pollution in Madrid (Spain). *Sci Total Environ*. (2004) 334:427–34. doi: 10.1016/j.scitotenv.2004.04.069

21. Pi T, Wu H, Li X. Does air pollution affect health and medical insurance cost in the elderly: an empirical evidence from China. *Sustainability*. (2019) 11:1526. doi: 10.3390/su11061526

22. Xia F, Xing J, Xu J, Pan X. The short-term impact of air pollution on medical expenditures: evidence from Beijing. *J Environ Econ Manag*. (2022) 114:102680. doi: 10.1016/j.jeem.2022.102680

23. Liao L, Du M, Chen Z. Air pollution, health care use and medical costs: evidence from China. *Energy Econ*. (2021) 95:105132. doi: 10.1016/j.eneco.2021.105132

24. Jütting JP. Do community-based health insurance schemes improve poor people's access to health care? Evidence from rural Senegal. *World Dev*. (2004) 32:273–88. doi: 10.1016/j.worlddev.2003.10.001

25. Anyanwu JC, Erhijakpor AE. Health expenditures and health outcomes in Africa. *Afr Dev Rev*. (2009) 21:400–33. doi: 10.1111/j.1467-8268.2009.00215.x

26. Li H, Lu J, Li B. Does pollution-intensive industrial agglomeration increase residents' health expenditure? *Sust Cities Soc*. (2020) 56:102092. doi: 10.1016/j.scs.2020.102092

27. Li X, Shen JJ, Lu J, Wang Y, Sun M, Li C, et al. Household catastrophic medical expenses in eastern China: determinants and policy implications. *BMC Health Serv Res*. (2013) 13:506. doi: 10.1186/1472-6963-13-506

28. Benvenuto M, Avram A, Sambati FV, Avram M, Viola C. The impact of internet usage and knowledge-intensive activities on households' healthcare expenditures. *Int J Environ Res Public Health*. (2020) 17:4470. doi: 10.3390/ijerph17124470

29. Beck T, Levine R, Levkov A. Big bad banks? The winners and losers from bank deregulation in the United States. *J Fin*. (2010) 65:1637–67. doi: 10.1111/j.1540-6261.2010.01589.x

30. Lyu Y, Lu Y, Wu S, Wang Y. The effect of the belt and road initiative on firms' OFDI: evidence from China's greenfield investment. *Econ Res J*. (2019) 54:187–202.

31. Heyman F, Sjöholm F, Tingvall PG. Is there really a foreign ownership wage premium? Evidence from matched employer–employee data. *J Int Econ*. (2007) 73:355–76. doi: 10.1016/j.jinteco.2007.04.003

32. Rao H, Chen D, Shen F, Shen Y. Can green bonds stimulate green innovation in enterprises? *Evid China. Sust*. (2022) 14:15631. doi: 10.3390/su142315631

33. Wu Z, Xiao Y, Zhang J. Labor mobility and corporate investment—evidence from a Quasi-natural experiment in China. *Int Rev Econ Fin*. (2022) 80:1110–29. doi: 10.1016/j.iref.2022.04.001

34. Deschenes O, Greenstone M, Shapiro JS. Defensive investments and the demand for air quality: evidence from the NOx budget program. *Am Econ Rev*. (2017) 107:2958–89. doi: 10.1257/aer.20131002

35. Xu B-C, Xu X-N, Zhao J-C, Zhang M. Influence of internet use on commercial health insurance of Chinese residents. *Front Public Health*. (2022) 10:907124. doi: 10.3389/fpubh.2022.907124

36. Mellor D, Firth L, Moore K. Can the internet improve the wellbeing of the elderly? *Ageing Int*. (2008) 32:25–42. doi: 10.1007/s12126-008-9006-3

37. Erickson J, Johnson GM. Internet use and psychological wellness during late adulthood. *Can J Aging*. (2011) 30:197–209. doi: 10.1017/S0714980811000109

38. Cotten SR, Ford G, Ford S, Hale TM. Internet use and depression among retired older adults in the United States: a longitudinal analysis. *J Gerontol Ser B Psychol Sci Soc Sci*. (2014) 69:763–71. doi: 10.1093/geronb/gbu018

39. Heo J, Chun S, Lee S, Lee KH, Kim J. Internet use and wellbeing in older adults. *Cyberpsychol Behav Soc Netw*. (2015) 18:268–72. doi: 10.1089/cyber.2014.0549

40. Hong YA, Zhou Z, Fang Y, Shi L. The digital divide and health disparities in China: evidence from a national survey and policy implications. *J Med Internet Res*. (2017) 19:e7786. doi: 10.2196/jmir.7786

41. Chen Y, Zhang L, Wei M. How does smart healthcare service affect resident health in the digital age? Empirical evidence from 105 cities of China. *Front Public Health*. (2021) 9:833687. doi: 10.3389/fpubh.2021.833687

42. Alam S, Shuaib M, Ahmad S, Jayakody DNK, Muthanna A, Bharany S, et al. Blockchain-based solutions supporting reliable healthcare for fog computing and internet of medical things (IoMT) integration. *Sustainability*. (2022) 14:15312. doi: 10.3390/su142215312

43. Wu H, Ba N, Ren S, Xu L, Chai J, Irfan M, et al. The impact of internet development on the health of Chinese residents: transmission mechanisms and empirical tests. *Socio Econ Plann Sci*. (2022) 81:101178. doi: 10.1016/j.seps.2021.101178

44. Lv Q, Jiang Y, Qi J, Zhang Y, Zhang X, Fang L, et al. Using mobile apps for health management: a new health care mode in China. JMIR mHealth and uHealth. (2019) 7:e10299. doi: 10.2196/10299

45. Chen L, Liu W. The effect of Internet access on body weight: evidence from China. *J Health Econ*. (2022) 85:102670. doi: 10.1016/j.jhealeco.2022.102670

frontiers | Frontiers in Public Health

TYPE Review
PUBLISHED 04 July 2023
DOI 10.3389/fpubh.2023.1102185

OPEN ACCESS

EDITED BY
Grant Murewanhema,
University of Zimbabwe, Zimbabwe

REVIEWED BY
Ricardo Valentim,
Federal University of Rio Grande do Norte, Brazil
Shailesh Tripathi,
Tampere University of Technology, Finland

*CORRESPONDENCE
Anelisa Jaca
✉ anelisa.jaca@mrc.ac.za

†These authors have contributed equally to this work and share first authorship

RECEIVED 18 November 2022
ACCEPTED 19 June 2023
PUBLISHED 04 July 2023

CITATION
Okeibunor JC, Jaca A, Iwu-Jaja CJ, Idemili-Aronu N, Ba H, Zantsi ZP, Ndlambe AM, Mavundza E, Muneene D, Wiysonge CS and Makubalo L (2023) The use of artificial intelligence for delivery of essential health services across WHO regions: a scoping review. *Front. Public Health* 11:1102185. doi: 10.3389/fpubh.2023.1102185

The use of artificial intelligence for delivery of essential health services across WHO regions: a scoping review

Joseph Chukwudi Okeibunor[1†], Anelisa Jaca[2*†], Chinwe Juliana Iwu-Jaja[2], Ngozi Idemili-Aronu[3], Housseynou Ba[1], Zukiswa Pamela Zantsi[2], Asiphe Mavis Ndlambe[2], Edison Mavundza[1], Derrick Muneene[4], Charles Shey Wiysonge[2,5] and Lindiwe Makubalo[1]

[1]World Health Organization Regional Office for Africa, Brazzaville, Republic of Congo, [2]Cochrane South Africa, South African Medical Research Council, Cape Town, South Africa, [3]Department of Sociology/Anthropology, University of Nigeria, Nsukka, Nigeria, [4]World Health Organization, Geneva, Switzerland, [5]HIV and Other Infectious Diseases Research Unit, South African Medical Research Council, Durban, South Africa

Background: Artificial intelligence (AI) is a broad outlet of computer science aimed at constructing machines capable of simulating and performing tasks usually done by human beings. The aim of this scoping review is to map existing evidence on the use of AI in the delivery of medical care.

Methods: We searched PubMed and Scopus in March 2022, screened identified records for eligibility, assessed full texts of potentially eligible publications, and extracted data from included studies in duplicate, resolving differences through discussion, arbitration, and consensus. We then conducted a narrative synthesis of extracted data.

Results: Several AI methods have been used to detect, diagnose, classify, manage, treat, and monitor the prognosis of various health issues. These AI models have been used in various health conditions, including communicable diseases, non-communicable diseases, and mental health.

Conclusions: Presently available evidence shows that AI models, predominantly deep learning, and machine learning, can significantly advance medical care delivery regarding the detection, diagnosis, management, and monitoring the prognosis of different illnesses.

KEYWORDS

artificial intelligence, deep learning, machine learning, non-communicable diseases, communicable diseases artificial intelligence, communicable diseases

1. Introduction

Artificial intelligence (AI) refers to the simulation of intellectual human behavior by computers. AI can be designed using lots of algorithms including machine learning (ML), deep learning (DL), natural language processing (NLP), support vector machine (SVM), and the artificial neural network (ANN) (1). These algorithms assist the system to identify the expected response which informs the computer what to expect (2). ML is the technique used in precision medicine for predicting treatment procedures and disease outcomes in patients (3). On the other hand, DL is a form of AI technique that is used health care to

identify potential cancerous cells using in radiology images beyond what can be perceived by the human eye. This method can promote faster learning without being prompted (3). Another form of AI, i.e., NLP is related to the use of software programming to understand and manipulate natural language text or speech for practical purposes (4). This involves dealing with large volumes of clinical data and health literacy in the health sector (5). SVM is an algorithm used to assemble a classification system for model classification and trend. The ANN model is used to comprehend the reasoning and functioning of connection between neurons (1). ANN has been used to solve different issues by building mathematical models that imitate natural activities of the brain (1).

AI has several advantages, i.e., it is reliable, cost-effective, solves complex issues, and limits data loss (6). AI is applied in fields including business, engineering, or medical care. In medical care, this technology is used for diagnosis, therapy, and prognosis (7). AI is a rapidly evolving field in medical care, with great potential to inform evidence-based decision making and ultimately improve health outcomes. It has been applied across various fields including robotics, medical diagnosis, medical statistics, and human biology (8). This technology plays a role in addressing certain issues within the health system which comprise staff shortages, poor administration of health services (e.g., billing, repayments, and insurance fraud exposure), and poor infrastructure; to support the delivery of high-quality healthcare (4, 7). AI also has the potential to impact on several aspects, including clinical decision at points of care, drug research, and disease predictions, amongst others. This has been said to improve efficiency, safety, and access to medical care services (2, 6, 8, 9).

Therefore, the AI technology is necessary to help manage of medical care services, to make decisions concerning disease prediction, diagnosing and treatment plans for patients (10). The current challenges (e.g., difficulty accessing health facilities in time, poor quality of health care, staff shortages) within the health system of low- and middle-income countries (LMICs) warrant the implementation and use of this technology (10). It is likely that this technology is predominantly applied in high-income countries (HICs) as LMICs may not have the infrastructure for the technology in their healthcare systems (10). AI can be used to manage various diseases, namely, diabetes, cancer, emerging infectious diseases, sexually transmitted diseases, and mental health illnesses. This technique has been utilized to predict risk and diagnosis of diabetes predicated on genomic and EHR data, respectively (11). It has also been used to predict risk of complications such as nephropathy and retinopathy (11). In cancer, AI can be used to analyze imaging data obtained during routine cancer care, i.e., disease classification, detection, segmentation, characterization, and monitoring. This saves time and helps radiologists achieve better outcomes and identify cancerous lesions that could be missed by humans (12). Furthermore, AI models are also useful in predicting the progression of disease and mortality in patients infected with emerging infectious diseases, namely, the severe acute respiratory syndrome (SARS), H1N1 influenza virus, Middle East respiratory syndrome coronavirus (MERS-CoV) as it has been recently done in SARS-CoV2 (13). Additionally, this method has been widely applied and has been used as an intervention for mental health issues. AI has been reported to be effective in managing mental health issues, i.e., reducing anxiety through detecting emotional changes and thought patterns and increasing thinking styles (14).

AI applications in mental health can bring insights into new treatment approaches. This technique has also been used to predict the diagnosis of sexually transmitted infections including HIV, as these are global public health concerns (15). This method has also been used in HIV prevention such as identifying potential PrEP candidates at risk of infection in Kenya and Uganda (16). One study evaluated the performance of AI in predicting HIV, syphilis, gonorrhea, chlamydia in an Australian cohort among men who have sex with men and reported that this technique is accurate (17). This type of research has mainly been conducted in Belgium, China, Italy, Korea, Turkey, and USA while there needs to be research conducted in LMICs, specifically in Africa (11, 13–16).

While health professionals in HICs may have the expertise to use the AI techniques, there may be a serious need to build capacity around this technology among professionals in LMICs (10). This implies that this technology would not work in LMICs where health care practitioners do not have the capacity to apply AI and interpret AI results. It is, however, necessary to assess where and for what conditions AI is being used around the world, thus the need for this scoping review. The objective of the review was to map out and synthesize the available evidence on the use of AI to deliver medical care services, globally and regionally.

2. Materials and methods

We conducted a scoping review as per the methodology defined by Arskey and O'Malley (18). A scoping review is a methodology that is used to chart key concepts and evidence available in a particular field. The field of AI is rapidly developing hence we used this methodology to undertake this review.

2.1. Search strategy

Two authors (Anelisa Jaca and Chinwe Juliana Iwu-Jaja) conducted a search in PubMed on 07 March 2022 and Scopus on 16 March 2022. The following combination of key words was used for the search: ("Artificial intelligence" OR AI OR "machine learning" OR "machine intelligence" OR "deep learning") AND ("health care" OR health OR "health delivery"). No language or date restrictions were employed. We first developed and implemented a search strategy in PubMed, which was afterwards adapted for Scopus.

2.2. Study selection

Titles and abstracts of identified records were independently screened by two researchers, Zukiswa Pamela Zantsi (ZPZ) and Asiphe Mavis Ndlambe (AMN), to identify potentially eligible records. Abstracts of records judged to be potentially eligible by one or both researchers were re-screened by a second pair of more experienced researchers, Anelisa Jaca (AJ) and Chinwe Juliana Iwu-Jaja (CJI). The latter made the decision on potentially eligible

studies through discussion and consensus. AJ and CJI then assessed the full text of potentially eligible studies and included publications of primary studies which: reported on the use of AI, addressed a health condition, assessed the effectiveness of the AI method used, and were published in a peer-reviewed journal in English. We excluded reviews.

2.3. Data extraction and analysis

A piloted data extraction form containing a list of data of interest and their definitions was used to extract data from eligible studies. Data were extracted independently by AJ, ZPZ, CJI, AMN, and Edison Mavundza (EM). The extracted data included the first author's name, year of publication, study population, country where the study was conducted, aim of the study, health issue, AI method, application of AI, and findings. The WHO region and category of health issue were also charted. We used a narrative synthesis method to analyse and report the key concepts and findings related to AI applications on medical care delivery. We did not evaluate the methodological quality of included studies since our aim was to identify and map the available evidence on the use of AI to deliver essential medical services. Three authors (AJ, CJI, and Charles Shey Wiysonge) had weekly meetings to discuss progress, findings, and next steps.

3. Results

3.1. Search results

Figure 1 shows the search and selection process for the scoping review. The literature search produced a total of 172,375 articles, including 11,695 from PubMed and 160,680 from Scopus. The first pair of researchers (AMN and ZPZ) screened these records and considered 1,129 publications to be potentially eligible for inclusion in the scoping review. The second pair (AJ and CJI) reviewed the 1,129 abstracts and found 801 to be potentially eligible for the review. Of the 801 potentially eligible articles, we randomly selected 100 publications whose full texts we obtained and assessed for eligibility. During the random selection, we used a systematic approach where we counted from the first article that appeared on Mendeley and selected every seventh publication. A systematic review involving all 801 potentially eligible studies is currently being undertaken. Of the 100 potentially eligible studies selected for assessment, 91 met inclusion criteria. Figure 1 shows the search and selection process.

3.2. WHO Regions

The characteristics of included studies are reported in detail in Supplementary Table 1.

Each of the six WHO regions had at least one publication included in the scoping review. Most of the included studies were conducted in WHO Region of the Americas (AMR) and European Region (EUR). The AMR had 28 studies (30.8%), which were conducted in the United States of America (Arizona, Florida, New York, Utah, Maryland, and Massachusetts), Mexico, Brazil, Chile, and Canada (19–46). EUR had 28 studies (30.8%) from France, the United Kingdom, Switzerland, Spain, Sweden, Turkey, Italy, and Germany (47–73). The Eastern Mediterranean Region (EMR) represented by Saudi Arabia, Iran, Iraq, and Pakistan had 14 studies (15.4%) (74–87). The South-East Asian Region (SEAR), represented by India, had 4 publications (4.3%) (88–91). The Western Pacific Region (WPR) represented by Australia, China and South Korea, had 11 publications (12.1%) (92–102). There was only one publication from the African Region (AFR), from Nigeria (1.1%) (103). Two publications included two countries each; one had China in WPR and UK in EUR and WPRO (2.2%) (104), while another involved Brazil in AMR and India in SEAR (105). Two of the publications (2.2%) were global studies and one study was conducted in Taiwan which does not fall under any WHO region (1.1%) (106–108).

3.3. Artificial intelligence methods and applications

3.3.1. Single approaches

a. **Artificial intelligence:** Eight publications reported the AI broad technique in the use of developing therapy as the intervention for infectious diseases, for diagnosing COVID-19 and mental health conditions, and as screening tools for diabetes and cancer (19, 36, 57, 62, 75, 97, 104, 109). This technique was also used in designing vaccination, measuring and increasing medication adherence in non-communicable diseases, identifying genomic sequences, and developing drugs and vaccines for COVID-19 (19, 36, 57, 62, 75, 97, 104, 109).
b. **Machine learning:** Machine learning (ML) was reported in 46 studies to analyse, classify, diagnose, manage, monitor, and predict different health conditions or diseases (e.g., frailty, back pain, ischemic stroke, cancer, COVID-19, tuberculosis, diabetes, mortality, hypertension, mental health conditions, bacterial vaginosis, and heart disease) (21, 22, 25–27, 30, 33, 37, 39, 41, 43, 46, 47, 50, 52, 54, 56, 58, 60, 63, 65, 67, 70, 71, 73, 76, 77, 81, 82, 84, 87, 92, 96, 101, 103, 105, 108, 110). This approach was also used to create patient re-admission files, pre-authorization in health insurance, and for finding missed cases of disease; these all form a significant part in delivering medical care services (30, 76, 111).
c. **Deep learning:** Fifteen studies reported the use of deep learning (DL) in detecting cardiovascular disease, predicting mortality and cancer, diagnosing asthma, classifying cancer subtypes, pre-screening for COVID-19, and analyzing diseases like macular oedema (20, 24, 44, 55, 61, 64, 79, 83, 91, 93, 94, 98, 99, 106).
d. **Convolutional neural network:** Only one study mentioned the use of convolutional neural network (CNN) to diagnose cardiac diseases (112).
e. **Artificial intelligence optical microscopic -based technology:** One study reported the use of artificial intelligence optical microscopic (AIOM)-based technology in reproductive health to quantitatively measure sperm concentration and motility as well as seminal pH (107).

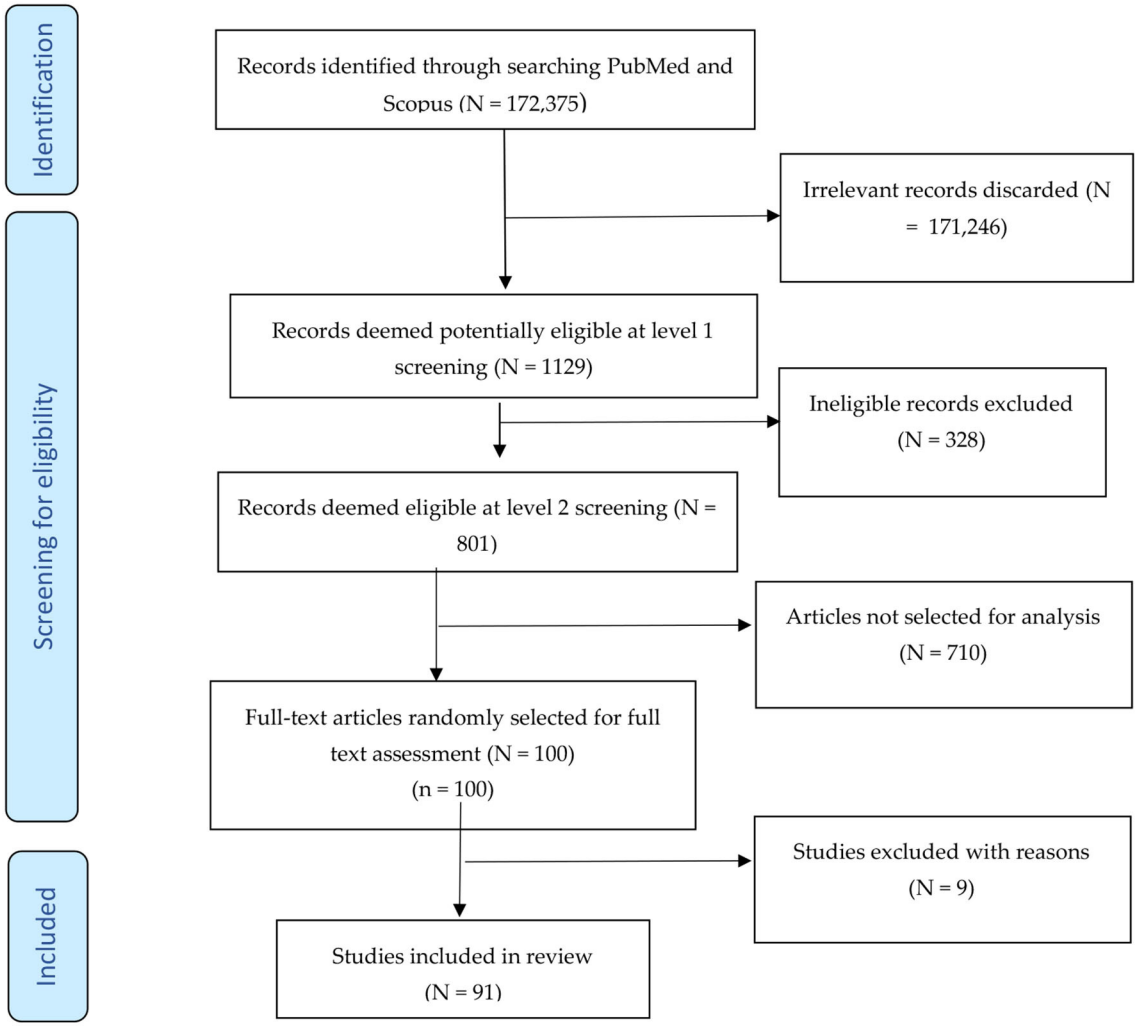


FIGURE 1
Study selection process for the scoping review.

f. **Artificial neural network:** Two studies investigated the use of artificial neural network (ANN) to predict infectious disease (COVID-19) and non-infectious disease (hearing loss) among noise-exposed workers (72, 78).
g. **Bayesian network:** The Bayesian network (BN) is defined as a graphical tool that can be employed to build models from data and or expert opinion. This method can be used to predict, detect, and diagnose disease (113). In one study, this method was used in predicting the prognosis of suicidal behavior (23).
h. **Deep neural network:** The deep neural network (DNN) was used in two studies to predict the mortality of patients in palliative care and classify breast cancer (29).
i. **Fuzzy K-means clustering algorithm:** The Fuzzy K-means clustering algorithm (FKCA) was used in one study to detect and classify cataract in normal, cataract, and post-cataract optical images (90).
j. **COVID Inception-ResNet model deep learning:** Almalki 2021 explored COVID Inception-ResNet (CoVIR-Net) model deep learning as a method for using chest X-rays to diagnose COVID-19 (40).

3.3.2. Combined approaches

Combining deep learning and machine learning with other approaches: Some of the included studies investigated the use of a combined AI approach for delivering medical care services. Three of those studies used DL and ML to diagnose and predict COVID-19, cardiovascular disease risk, hazardous drinkers, and the severity of alcohol-related problems (49, 102, 114). ML was also used in combination with artificial neural network to diagnose, predict and manage prognosis of nervous system disorders (69). One study reported the use of DL and NLP to quantify left and right ventricular dysfunction from electrocardiograms. DL was also used together with neural network mode to diagnose COVID-19 in chest X-ray images (68). In another study, DL was used with multi-head attention (MHA), Long Short-Term memory (LSM), and CNN (45).

3.4. Health conditions assessed using AI applications

A total of 21 studies focused on infectious diseases of various types (33, 40, 42, 45, 49, 52, 64, 65, 68, 72, 75, 80, 82, 83, 91, 100, 104, 106, 108, 114). Eighteen of these studies focused on COVID-19 while the remaining ones were on tuberculosis. Thirteen studies targeted cardiovascular diseases (including stroke, hypertension, ventricular dysfunction, and heart function) (31, 34, 39, 44, 46, 51, 59, 61, 71, 76, 87, 95, 112). These were mostly experimental studies conducted in WHO AMR, EMR, EUR, SEAR, and WPR regions for prediction and diagnostic purposes. Six studies focused on cancers, including prostate, lung, skin, and breast (48, 67, 73, 79, 92, 115). They were conducted in AFR, EUR, SEAR, and WPR. These investigations were mostly experimental studies for prediction and diagnostic purposes. There were 21 studies on conditions on assorted conditions ((22, 24, 26, 27, 30, 35, 36, 39, 41, 54, 55, 58, 60, 66, 69, 70, 78, 81, 85, 86, 98). These conditions include injuries, diet, sepsis, and drug overdose. Eight studies were on mental and cognitive health problems including various forms of depression and dementia (37, 50, 53, 57, 84, 101). These studies were conducted in EUR, AMR, and SEAR. The studies mostly focused on prediction and diagnosis. Only one study was on reproductive health where AI was used for diagnostic purposes in men (107).

4. Discussion

This paper highlights currently used AI techniques and applications. The AI techniques identified include machine learning, deep learning, convolutional neural network, artificial intelligence optical microscopic-based technology, artificial neural network, Bayesian network, deep neural network, Fuzzy K-means clustering algorithm, COVID Inception-ResNet model deep learning, natural language processing, neural network mode, and long short-term memory. The AI techniques were used for four main groups of medical care services, including: (i) detection and diagnosis; (ii) classification; (iii) treatment, support, and prognosis; and (iv) management of research and clinical care. Most of the studies focused on the use of machine learning and distance learning as applications to detect and diagnose different diseases. These include infectious diseases (COVID-19 and tuberculosis); cardiovascular diseases (ischemic stroke, cardiomyopathy, hypertension); metabolic diseases (diabetes); cancers (breast, prostate, diffuse gliomas, skin); and mental diseases (schizophrenia, dementia, suicidal behavior). These techniques have also been used in hospital and research administration. Generally, machine learning and deep learning show the possibility of being used to improve the competence of clinical and research procedures which will be beneficial to good health.

It is important to note that these investigations were predominantly conducted in the WHO Region of the Americas and the European Region than in other regions. We only found one study conducted in Africa in this sample of studies. The difference concerning distributions of publications in the WHO regions showed that there is a lack of research conducted in low and middle-income countries around this field of study. Most healthcare facilities with lack of resources infrastructure, specifically, low-and middle-income countries (LMICs), do not have digital infrastructure to implement AI in their settings. HICs on the other hand, with highly skilled healthcare workers who can explain AI results regarding clinical scenario while in LMICs, all this may be lacking. It is important to note that for AI to be fully functional, it first must be available, accessible, and sustainable (10). A good example is where AI is applied in radiology, where its functionality involves services like imaging hardware, servers, information technology, quick internet service, picture storage and communication system, electronic medical records, and cloud services. This shows that it would be challenging to establish AI in low resource settings where is lack of experts to interpret its outputs and to apply them appropriately (10). Therefore, if AI would be implemented successfully in LMICs, healthcare professionals would need to be educated and trained on how to use the technology. AI implementation involves a lot of processes for it to perform well and important aspect is using the same data from the same source as the training set. Currently, most data to develop AI come from HICs, with some from middle-income countries (116). Health experts recommend that LMICs regulate AI processes together with global health organizations who would give them support (116). This would help healthcare workers in LMICs with successfully implementing and applying this technology. Other benefits of applying AI in healthcare in LMICs would be improving the state of health systems and decreasing medical costs such as those of screening (117). Furthermore, costs related to treatment plans that need expensive tools and specialized expertise especially in rural and remote settings (117).

There is, however, a relatively high number of publications in the Eastern Mediterranean Region, mainly involving studies conducted in Saudi Arabia. The lack of studies from under resourced regions may also be due to insufficient resources and AI knowledge among healthcare practitioners, especially in the WHO African Region. In view of the above, AI can be greatly beneficial to healthcare services in LMICs although its introduction would be quite different from what is done in HICs (10). For this scoping review, we used a random sample of currently available peer-reviewed publications to map out and synthesize the available evidence on the use of AI to deliver medical care services globally and regionally. The review suggests that AI is predominantly applied in high-income countries, with its use still emerging in low-resource settings (such as the WHO African Region) perhaps because health institutions in these settings do not have the infrastructure for use of this technology. This scoping review suggests that there is value in undertaking a systematic review and will proceed with conducting the review on the topic. The systematic review will focus on discussing the use and effectiveness of AI in delivering healthcare services. Furthermore, the review will discuss and compare applications including diagnosis and treatment of disease, patient engagement and adherence, and administrative services.

Linear models, including linear regression, multiple regression and multivariate linear regression models have also been used in medical research. Linear and multiple regression models are methods used to predict and assess interactions between the different datasets (118). The linear regression model is also used

to address different research questions and study aims (118). Multiple linear regression has been used to predict the length of stay for patients undergoing treatment heart disease, diabetes, hypertension, cancer, and laparoscopic appendectomy by (119). Multivariate Regression Analysis of Variance) is a multiple test that combines all the tests on the significance of the single regression coefficients (120).

5. Strengths and limitations of the study

The present review achieved its aim of mapping out and synthesizing literature on the use of AI methods in medical care. However, it is important to note that a broader review, i.e., a systematic review, would further illuminate the gaps in literature. A limitation of this study is that it did not evaluate the effectiveness of the different AI techniques in medica care services. A research question on the effectiveness of AI techniques within the different health issues require that another review be conducted. In review of the above, it must be considered that conducting rapid rather than systematic reviews to address effectiveness questions would be beneficial since there are constant publications around this field. A rapid review would ensure that relevant evidence is collected and disseminated in time.

6. Conclusions

Currently available evidence shows that the AI techniques are commonly used to deliver medical care services, especially in HICs. The commonly used methods for the detection, diagnosis, management, and monitoring the prognosis of different diseases are deep learning, and machine learning. The use of AI in the various health issues, namely, infectious diseases (COVID-19 and TB); metabolic diseases (diabetes); cardiovascular disease (ischemic stroke, cardiomyopathy, hypertension), cancers (breast, prostate, diffuse gliomas, skin) and mental diseases (schizophrenia, dementia, suicidal behavior) has shown positive outcomes. Other conditions in which the application of AI has shown positivity, include, frailty, low back pain, oral leucoplakia, open wound mortality, pressure injuries, primary progressive aphasia, dementia, lung function, asthma, and growth hormone deficiency. Further research is required in the use of other AI techniques to advance medical care delivery, especially in WHO African, Eastern Mediterranean, South-East Asian, and Western Pacific regions. AI methods are becoming more available for researchers and clinicians to apply, and it is probable that this field will continue to grow.

Author contributions

AJ designed the search strategy with an important input from CI-J and CW. AJ and CI-J conducted literature searches. AN and ZZ screened the search output and AJ and CI-J re-screened the articles. AN, ZZ, AJ, CI-J, and EM extracted data from eligible articles. AJ, CI-J, and JO wrote the first draft of the manuscript. CW, JO, DM, LM, NI-A, and HB guided the project and critically revised the intellectual content of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the World Health Organization (WHO).

Acknowledgments

We would like to acknowledge the South African Medical Research Council for availing the premises and equipment to undertake this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1102185/full#supplementary-material

References

1. Alsedrah MK. *Running Head : ARTIFICIAL INTELLIGENT Artificial Intelligence Advanced Analysis and Design : CNIT 380 Instructors : Dr . Hiba Tabbarah & Mr . Abdullah Abdulghafar Semester : Fall 2017. Section : U1 Mariam Khaled AlSedrah," no. December 2017* (2018)

2. Sunarti S, Fadzlul Rahman F, Naufal M, Risky M, Febriyanto K, Masnina R. Artificial intelligence in healthcare: opportunities and risk for future. *Gaceta Sanitaria.* (2021) 35:S67–S70. doi: 10.1016/j.gaceta.2020.12.019

3. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Jama*. (2019) 6:94–98. doi: 10.7861/futurehosp.6-2-94

4. Reddy S. "Use of artificial intelligence," In: *Healthcare Delivery* (2018).

5. Yuvaraj D, Uvaze AM, Sivaram M. "Materials today : proceedings a study on the role of natural language processing in the healthcare sector," In: *Materials Today: Proceedings* (2021).

6. Jiang F. *Artificial Intelligence in Healthcare : Past, Present and Future* (2017).

7. Reddy S. *Artificial intelligence and healthcare — why they need each other ?*, pp. 2020–2022 (2021).

8. Chen M, Decary M. Artificial intelligence in healthcare: an essential guide for health leaders. *Healthcare Manage Forum.* (2020) 33:10–8. doi: 10.1177/0840470419873123

9. Hosny A, Aerts HJWL, Oncology R. HHS public access. *Jama.* (2021) 366:955–56. doi: 10.1126/science.aay5189

10. Mollura DJ, Culp MP, Pollack E, Battino G, Scheel JR, Mango VL, et al. Artificial intelligence in low- and middle-income countries: innovating global health radiology. *Radiology.* (2020) 297:513–20. doi: 10.1148/radiol.2020201434

11. Singla R, Singla A, Gupta Y, Kalra S. *Artificial Intelligence/Machine Learning in Diabetes Care*. (2019). p. 495–7.

12. Dacoregio MI, Batalini F, Moraes FY. *An Overview of Artificial Intelligence in Oncology* (2022).

13. Chiu H-YR, Hwang C-K, Chen S-Y, Shih F-Y, Han H-C, King C-C, et al. Machine learning for emerging infectious disease field responses. *Scientific Reports.* (2022) 22:1–13. doi: 10.1038/s41598-021-03687-w

14. Zhou S, Zhao J, Zhang L. Application of artificial intelligence on psychological interventions and diagnosis: an overview. *Dell.* (2022) 13:1–7. doi: 10.3389/fpsyt.2022.811665

15. Xu X, Ge Z, Chow EPF, Yu Z, Lee D, Wu J, et al. *A Machine-Learning-Based Risk-Prediction Tool for HIV and Sexually Transmitted Infections Acquisition over the Next 12 Months*. (2022). p. 2016–2021.

16. Marcus JL, Sewell WC. Emerging approaches to ending the epidemic. *Jama.* (2021) 17:171–9. doi: 10.1007/s11904-020-00490-6

17. Bao Y, Medland NA, Fairley CK, Wu J, Shang X, Chow EPF, et al. Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. *Jama.* (2021) 82:48–59. doi: 10.1016/j.jinf.2020.11.007

18. Arksey H, Malley LO, Arksey H, Malley LO. *Scoping studies : towards a methodological framework Scoping Studies : Towards a Methodological Framework*. (2014). p. 37–41.

19. Arenas-Cavalli JT, Abarca I, Rojas-Contreras M, Bernuy F, Donoso R. Correction: clinical validation of an artificial intelligence-based diabetic retinopathy screening tool for a national health system. *Eye.* (2021) 35:2910. doi: 10.1038/s41433-021-01690-z

20. Adedinsewo DA, Johnson PW, Douglass EJ, Attia IZ, Phillips SD, Goswami RM, et al. Detecting cardiomyopathies in pregnancy and the postpartum period with an electrocardiogram-based deep learning model. *Euro Heart J Digit Health.* (2021) 2:586–96. doi: 10.1093/ehjdh/ztab078

21. Arefeen MA, Nimi ST, Rahman MS, Arshad SH, Holloway JW, Rezwan FI. Prediction of lung function in adolescence using epigenetic aging: a machine learning approach. *Methods Protocols.* (2020) 3:24. doi: 10.3390/mps3040077

22. Bedoya AD, Futoma J, Clement ME, Corey K, Brajer N, Lin A, et al. Machine learning for early detection of sepsis: an internal and temporal validation study. *JAMIA open.* (2020) 3:252–60. doi: 10.1093/jamiaopen/ooaa006

23. Barros J, Morales S, García A, Echávarri O, Fischman R, Szmulewicz M, et al. Recognizing states of psychological vulnerability to suicidal behavior: a Bayesian network of artificial intelligence applied to a clinical sample. *BMC psychiatry.* (2020) 20:138. doi: 10.1186/s12888-020-02535-x

24. Akkus Z, Cai J, Boonrod A, Zeinoddini A, Weston AD, Philbrick KA, et al. A survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow. *J Am Coll Radiol : JACR.* (2019) 16:1318–28. doi: 10.1016/j.jacr.2019.06.004

25. Banda JM, Sarraju A, Abbasi F, Parizo J, Pariani M, Ison H, et al. Finding missed cases of familial hypercholesterolemia in health systems using machine learning. *NPJ Digit Med.* (2019) 2:23. doi: 10.1038/s41746-019-0101-5

26. Badger J, LaRose E, Mayer J, Bashiri F, Page D, Peissig P. Machine learning for phenotyping opioid overdose events. *J Biomed Informat.* (2019) 94:103185. doi: 10.1016/j.jbi.2019.103185

27. Barton C, Chettipally U, Zhou Y, Jiang Z, Lynn-Palevsky A, Le S, et al. Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs. *Comput Biol Med.* (2019) 109:79–84. doi: 10.1016/j.compbiomed.2019.04.027

28. Anan T, Kajiki S, Oka H, Fujii T, Kawamata K, Mori K, et al. Effects of an artificial intelligence-assisted health program on workers with neck/shoulder pain/stiffness and low back pain: randomized controlled trial. *JMIR mHealth and uHealth.* (2021) 9:e27535. doi: 10.2196/27535

29. Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. *BMC Med Informat Decis Mak.* (2018) 18:122. doi: 10.1186/s12911-018-0677-8

30. Araújo FHD, Santana AM, de A Santos Neto P. Using machine learning to support healthcare professionals in making preauthorisation decisions. *International J Med informat.* (2016) 94:1–7. doi: 10.1016/j.ijmedinf.2016.06.007

31. Alanazi EM, Abdou A, Luo J. Predicting risk of stroke from lab tests using machine learning algorithms: development and evaluation of prediction models. *JMIR Format Res.* (2021) 5:e23440. doi: 10.2196/23440

32. Beck D, Foster JA. Machine learning techniques accurately classify microbial communities by bacterial vaginosis characteristics. *PloS one.* (2014) 9:e87830. doi: 10.1371/journal.pone.0087830

33. Al Bulushi Y, Saint-Martin C, Muthukrishnan N, Maleki F, Reinhold C, Forghani R. Radiomics and machine learning for the diagnosis of pediatric cervical non-tuberculous mycobacterial lymphadenitis. *Scientific Reports.* (2022) 22:585. doi: 10.1038/s41598-022-06884-3

34. Vaid A, Johnson KW, Badgeley MA, Somani SS, Bicak M, Landi I, et al. Using deep-learning algorithms to simultaneously identify right and left ventricular dysfunction from the electrocardiogram. *JACC: Cardiovascular Imag.* (2022) 22:895. doi: 10.1016/j.jcmg.2021.08.004

35. Killian MO, Payrovnaziri SN, Gupta D, Desai D, He Z. Machine learning-based prediction of health outcomes in pediatric organ transplantation recipients. *JAMIA open.* (2021) 4:ooab008. doi: 10.1093/jamiaopen/ooab008

36. Rocha TAH, de Almeida DG, Kozhumam AS, da Silva NC, Thomaz EBAF, Queiroz RCd, et al. Microplanning for designing vaccination campaigns in low-resource settings : a geospatial artificial intelligence-based framework. *Vaccine.* (2021) 39:6276–82. doi: 10.1016/j.vaccine.2021.09.018

37. Aschwanden D, Aichele S, Ghisletta P, Terracciano A, Kliegel M, Sutin AR, et al. Predicting cognitive impairment and dementia: a machine learning approach. *J Alzheimer's Dis JAD.* (2020) 75:717–28. doi: 10.3233/JAD-190967

38. Model L. HHS. *Public Access.* (2018) 27:461–8.

39. Alderden J, Pepper GA, Wilson A, Whitney JD, Richardson S, Butcher R, et al. Predicting pressure injury in critical care patients: a machine-learning model. *Am J Critic Care : an Offic Publicat Am Assoc Critic Care Nurses.* (2018) 27:461–8. doi: 10.4037/ajcc2018525

40. Almalki YE, Qayyum A, Irfan M, Haider N, Glowacz A, Alshehri FM, et al. A novel method for COVID-19 diagnosis using artificial intelligence in chest x-ray images. *Healthcare.* (2021) 9:5. doi: 10.3390/healthcare9050522

41. Anderson C, Bekele Z, Qiu Y, Tschannen D, Dinov ID. Modeling and prediction of pressure injury in hospitalized patients using artificial intelligence. *BMC Med Informat Decis mak.* (2021) 21:253. doi: 10.1186/s12911-021-01608-5

42. Alves MA, Castro GZ, Oliveira BAS, Ferreira LA, Ramírez JA, Silva R, et al. Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs. *Comput Biol Med.* (2021) 132:104335. doi: 10.1016/j.compbiomed.2021.104335

43. Akiki RK, Anand RS, Borrelli M, Sarkar IN, Liu PY, Chen ES. Predicting open wound mortality in the ICU using machine learning. *J Emerg Critical Care Med.* (2021) 5:13. doi: 10.21037/jeccm-20-154

44. Kshatriya BSA, Sagheb E, Wi C-I, Yoon J, Seol HY, Juhn Y, et al. Identification of asthma control factor in clinical notes using a hybrid deep learning model. *BMC Med Informat decision Mak.* (2021) 21:272. doi: 10.1186/s12911-021-01633-4

45. Abbasimehr H, Paki R. Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos Soliton Fractal.* (2021) 142:110511. doi: 10.1016/j.chaos.2020.110511

46. Amaratunga D, Cabrera J, Sargsyan D, Kostis JB, Zinonos S, Kostis WJ. Uses and opportunities for machine learning in hypertension research. *Int J Cardiol Hypertens.* (2020) 5:100027. doi: 10.1016/j.ijchy.2020.100027

47. Awada H, Gurnari C, Durmaz A, Awada H, Pagliuca S, Visconte V. Personalized risk schemes and machine learning to empower genomic prognostication models in myelodysplastic syndromes. *Int J Mol Sci.* (2022) 5:100027. doi: 10.3390/ijms23052802

48. Bechelli S, Delhommelle J. Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. *Bioengineering.* (2022) 22:65. doi: 10.3390/bioengineering9030097

49. Ancochea J, Izquierdo JL, Soriano JB. Evidence of gender differences in the diagnosis and management of coronavirus disease patients: an analysis of electronic health records using natural language processing and machine learning. *J Women's Health.* (2002) 30:393–404. doi: 10.1089/jwh.2020.8721

50. Andersson S, Bathula DR, Iliadis SI, Walter M, Skalkidou A. Predicting women with depressive symptoms postpartum with machine learning methods. *Scientific Rep.* (2021) 11:7877. doi: 10.1038/s41598-021-86368-y

51. Attia ZI, Harmon DM, Behr ER, Friedman PA. Application of artificial intelligence to the electrocardiogram. *European heart J.* (2021) 42:4717–30. doi: 10.1093/eurheartj/ehab649

52. Badimon L, Robinson EL, Jusic A, Carpusca I, deWindt LJ, Emanueli C, et al. Cardiovascular RNA markers and artificial intelligence may improve COVID-19 outcome: a position paper from the EU-CardioRNA COST Action CA17129. *Cardiovascul Res.* (2021) 117:1823–40. doi: 10.1093/cvr/cvab094

53. Balea-Fernandez FJ, Martinez-Vega B, Ortega S, Fabelo H, Leon R, Callico GM, et al. Analysis of risk factors in dementia through machine learning. *J Alzheimer's Dis JAD.* (2021) 79:845–61. doi: 10.3233/JAD-200955

54. Baron JM, Huang R, McEvoy D, Dighe AS. Use of machine learning to predict clinical decision support compliance, reduce alert burden, and evaluate duplicate laboratory test ordering alerts. *JAMIA Open.* (2021) 4:ooab006. doi: 10.1093/jamiaopen/ooab006

55. M. Araujo, van Dommelen P, Koledova E, Srivastava J. Using deep learning for individual-level predictions of adherence with growth hormone therapy. *Studies Health Technol Informat.* (2021) 281:133–7. doi: 10.3233/SHTI210135

56. Abbas M, Somme D, Le Bouquin Jeannes R. "Machine learning-based physical activity tracking with a view to frailty analysis," In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference.* (2020). p. 3917–3920.

57. Barrera A, Gee C, Wood A, Gibson O, Bayley D, Geddes J. Introducing artificial intelligence in acute psychiatric inpatient care: qualitative study of its use to conduct nursing observations. *Evid Based Ment Health.* (2020) 23:34–8. doi: 10.1136/ebmental-2019-300136

58. Battineni G, Sagaro GG, Chinatalapudi N, Amenta F. Applications of machine learning predictive models in the chronic disease diagnosis. *J Personal Med.* (2020) 10:2. doi: 10.3390/jpm10020021

59. Bajaj R, Eggermont J, Grainger SJ, Räber L, Parasa R, Khan AHA, et al. Machine learning for atherosclerotic tissue component classification in combined near-infrared spectroscopy intravascular ultrasound imaging: validation against histology. *Atherosclerosis.* (2022) 22:234. doi: 10.1016/j.atherosclerosis.2022.01.01

60. Álvarez JD, Matias-Guiu JA, Cabrera-Martín MN, Risco-Martín JL, Ayala JL. An application of machine learning with feature selection to improve diagnosis and classification of neurodegenerative disorders. *BMC bioinformat.* (2019) 20:491. doi: 10.1186/s12859-019-3027-7

61. Ashfaq A, Sant'Anna A, Lingman M, Nowaczyk S. Readmission prediction using deep learning on electronic health records. *J Biomed Informatics.* (2019) 97:103256. doi: 10.1016/j.jbi.2019.103256

62. Baskaran V, Bali RK, Arochena H, Naguib RNG, Wallis M, Wheaton M. "Knowledge creation using artificial intelligence: a twin approach to improve breast screening attendance," In: *Conference proceedings : . Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference.* (2006). p. 4070–4073.

63. Andrew TW, Hamnett N, Roy I, Garioch J, Nobes J, Moncrieff MD. Machine-learning algorithm to predict multidisciplinary team treatment recommendations in the management of basal cell carcinoma. *British J Cancer.* (2022) 126:562–8. doi: 10.1038/s41416-021-01506-7

64. Altan G. DeepOCT: An explainable deep learning architecture to analyze macular edema on OCT images. *Eng Sci Technol Int J.* (2022) 126:562–8. doi: 10.1016/j.jestch.2021.101091

65. Banerjee A, Ray S, Vorselaars B, Kitson J, Mamalakis M, Weeks S, et al. Use of machine learning and artificial intelligence to predict SARS-CoV-2 infection from full blood counts in a population. *Int immunopharmacol.* (2020) 86:106705. doi: 10.1016/j.intimp.2020.106705

66. Arceo-Vilas A, Fernandez-Lozano C, Pita S, Pértega-Díaz S, Pazos A. Identification of predictive factors of the degree of adherence to the Mediterranean diet through machine-learning techniques. *PeerJ. Computer Sci.* (2020) 6:e287. doi: 10.7717/peerj-cs.287

67. Beinecke JM, Anders P, Schurrat T, Heider D, Luster M, Librizzi D, et al. Evaluation of machine learning strategies for imaging confirmed prostate cancer recurrence prediction on electronic health records. *Comput Biol Med.* (2022) 143:105263. doi: 10.1016/j.compbiomed.2022.105263

68. Bayram F, Eleyan A. COVID-19 detection on chest radiographs using feature fusion based deep learning. *Sign Image Video Process.* (2022) 22:836. doi: 10.1007/s11760-021-02098-8

69. Auger SD, Jacobs BM, Dobson R, Marshall CR, Noyce AJ. Big data, machine learning and artificial intelligence: a neurologist's guide. *Practic Neurol.* (2020) 21:4–11. doi: 10.1136/practneurol-2020-002688

70. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med informa Decis Mak.* (2020) 20:310. doi: 10.1186/s12911-020-01332-6

71. Abedi V, Avula V, Chaudhary D, Shahjouei S, Khan A, Griessenauer CJ, et al. Prediction of long-term stroke recurrence using machine learning models. *J Clinic Med.* (2021) 10:6. doi: 10.3390/jcm10061286

72. Abdulaal A, Patel A, Al-Hindawi A, Charani E, Alqahtani SA, Davies GW, et al. Clinical utility and functionality of an artificial intelligence-based app to predict mortality in COVID-19: mixed methods analysis. *JMIR Format Res.* (2021) 5:e27992. doi: 10.2196/27992

73. Adeoye J, Koohi-Moghadam M, Lo AWI, Tsang RK-Y, Chow VLY, Zheng L-W, et al. Deep learning predicts the malignant-transformation-free survival of oral potentially malignant disorders. *Cancers.* (2021) 13:23. doi: 10.3390/cancers13236054

74. Han JED, Liu X, Bunce C, Douiri A, Vale L, Blandford A, et al. Teleophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease (HERMES): a Cluster Randomised Superiority Trial with a linked Diagnostic Accuracy Study-HERMES study report 1-study protocol. *BMJ open.* (2022) 12:e055845. doi: 10.1136/bmjopen-2021-055845

75. Abubaker Bagabir S, Ibrahim NK, Abubaker Bagabir H, Hashem Ateeq R. Covid-19 and Artificial Intelligence: genome sequencing, drug development and vaccine discovery. *J Infect Public Health.* (2022) 15:289–96. doi: 10.1016/j.jiph.2022.01.011

76. Alzeer AH, Althemery A, Alsaawi F, Albalawi M, Alharbi A, Alzahrani S, et al. Using machine learning to reduce unnecessary rehospitalization of cardiovascular patients in Saudi Arabia. *Int J Med Informat.* (2021) 154:104565. doi: 10.1016/j.ijmedinf.2021.104565

77. Alloghani M, Aljaaf A, Hussain A, Baker T, Mustafina J, Al-Jumeily D, Khalaf M.Implementation of machine learning algorithms to create diabetic patient re-admission profiles. *BMC Med Informat Decis Mak.* (2019) 19:253. doi: 10.1186/s12911-019-0990-x

78. Aliabadi M, Farhadian M, Darvishi E. Prediction of hearing loss among the noise-exposed workers in a steel factory using artificial intelligence approach. *Int Archiv Occupat Environ Health.* (2015) 88:779–787. doi: 10.1007/s00420-014-1004-z

79. Ali AM, Mohammed AA. Improving classification accuracy for prostate cancer using noise removal filter and deep learning technique. *Multimedia Tools Applicat.* (2022) 22:569. doi: 10.1007/s11042-022-12102-z

80. Alafif T, Tehame AM, Bajaba S, Barnawi A, Zia S. Machine and deep learning towards COVID-19 diagnosis and treatment: survey, challenges, and future directions. *Int J Environ Res Public Health.* (2021) 18:3. doi: 10.3390/ijerph18031117

81. Alhorishi N, Almeziny M, Alshammari R. Using machine learning to predict early preparation of pharmacy prescriptions at psmmc - a comparison of four machine learning algorithms. *Acta Informatica Medica : AIM : journal of the Society for Medical Informatics of Bosnia & Herzegovina : casopis Drustva za medicinsku informatiku BiH.* (2021) 29:21–5. doi: 10.5455/aim.2021.29.21-25

82. Ali MH, Khan DM, Jamal K, Ahmad Z, Manzoor S, Khan Z. Prediction of multidrug-resistant tuberculosis using machine learning algorithms in SWAT, Pakistan. *J Healthcare Eng.* (2021) 21:2567080. doi: 10.1155/2021/2567080

83. Alruwaili M, Shehab A, Abd El-Ghany S. COVID-19 diagnosis using an enhanced inception-resnetv2 deep learning model in cxr images. *J Healthcare Eng.* (2021) 21:6658058. doi: 10.1155/2021/6658058

84. AlShorman O, Masadeh M, Heyat MBB, Akhtar F, Almahasneh H, Ashraf GM, et al. Frontal lobe real-time EEG analysis using machine learning techniques for mental stress detection. *J Integrat Neurosci.* (2022) 21:20. doi: 10.31083/j.jin2101020

85. Abdollahi M, Ashouri S, Abedi M, Azadeh-Fard N, Parnianpour M, Khalaf K, et al. Using a motion sensor to categorize nonspecific low back pain patients: a machine learning approach. *Sensors.* (2020) 20:12. doi: 10.3390/s20123600

86. Althobaiti T, Katsigiannis S, Ramzan N. Triaxial accelerometer-based falls and activities of daily life detection using machine learning. *Sensors.* (2020) 20:113. doi: 10.3390/s20133777

87. Almazroi AA. Survival prediction among heart patients using machine learning techniques. *Mathematic Biosci Eng MBE.* (2022) 19:134–145. doi: 10.3934/mbe.2022007

88. PrayGod G, Blevins M, Woodd S, Rehman AM, Jeremiah K, Friis H, et al. A longitudinal study of systemic inflammation and recovery of lean body mass among malnourished HIV-infected adults starting antiretroviral therapy in Tanzania and Zambia. *Europ J Clinic Nutri.* (2016) 70:499–504. doi: 10.1038/ejcn.2015.221

89. Lyu W, Yuan B, Liu S, Simon JE, Wu Q. Assessment of lemon juice adulteration by targeted screening using LC-UV-MS and untargeted screening using UHPLC-QTOF/MS with machine learning. *Food Chemistr.* (2022) 25:465. doi: 10.1016/j.foodchem.2021.131424

90. Acharya RU, Yu W, Zhu K, Nayak J, Lim T-C, Chan JY. Identification of cataract and post-cataract surgery optical images using artificial intelligence techniques. *J Med Syst.* (2010) 34:619–28. doi: 10.1007/s10916-009-9275-8

91. Id MA, Khandoker AH. *Detection of COVID-19 in smartphone-based breathing recordings : A pre-screening deep learning tool.* (2022). p. 1–25.

92. Alsinglawi B, Alshari O, Alorjani M, Mubin O, Alnajjar F, Novoa M, et al. An explainable machine learning framework for lung cancer hospital length of stay prediction. *Scientific Rep.* (2022) 22:695. doi: 10.1038/s41598-021-04608-7

93. Balasubramaniyan S, Jeyakumar V, Nachimuthu DS. Panoramic tongue imaging and deep convolutional machine learning model for diabetes diagnosis in humans. *Scientific Rep.* (2022) 26:695. doi: 10.1038/s41598-021-03879-4

94. Kim T. *Pathological Images Using Deep Transfer Learning.* (2021). p. 1–14.

95. Barbieri S, Mehta S, Wu B, Bharat C, Poppe K, Jorm L, et al. Predicting cardiovascular risk from national administrative databases using a combined

survival analysis and deep learning approach. *Int J Epidemiol.* (2021) 9:256. doi: 10.1093/ije/dyab258

96. Alazzam MB, Mansour H, Alassery F, Almulihi A. Machine learning implementation of a diabetic patient monitoring system using interactive e-app. *Computat Intell Neurosci.* (2021) 21:5759184. doi: 10.1155/2021/5759184

97. Abdulla A, Wang B, Qian F, Kee T, Blasiak A, Ong YH, et al. Project IDentif.AI: harnessing artificial intelligence to rapidly optimize combination therapy development for infectious disease intervention. *Adv Therapeutic.* (2020) 20:34. doi: 10.1002/adtp.202000034

98. Adegbosin AE, Stantic B, Sun J. Efficacy of deep learning methods for predicting under-five mortality in 34 low-income and middle-income countries. *BMJ open.* (2020) 10:e034524. doi: 10.1136/bmjopen-2019-034524

99. Wang M, Wei Z, Jia M, Chen L, Ji H. Deep learning model for multi-classification of infectious diseases from unstructured electronic medical records *BMC Med Informat Deci Mak.* (2022) 22:56. doi: 10.1186/s12911-022-01776-y

100. Al-Antari MA, Hua C-H, Bang J, Lee S. Fast deep learning computer-aided diagnosis of COVID-19 based on digital chest x-ray images. *Appl intelligen.* (2020) 11:1–18. doi: 10.21203/rs.3.rs-36353/v2

101. Bae YJ, Shim M, Lee WH. Schizophrenia detection using machine learning approach from social media content. *Sensors.* (2021) 21:751. doi: 10.3390/s21175924

102. Kim S-Y, Park T, Kim K, Oh J, Park Y, Kim D-J. A deep learning algorithm to predict hazardous drinkers and the severity of alcohol-related problems using K-NHANES. *Front Psychiatry.* (2021) 12:684406. doi: 10.3389/fpsyt.2021.684406

103. Ameh Joseph, Abdullahi M, Junaidu SB, Hassan Ibrahim H, Chiroma H. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). *Intell Syst Applicat.* (2022) 22:656. doi: 10.1016/j.iswa.2022.200066

104. Bai X, Wang H, Ma L, Xu Y, Gan J, Fan Z, et al. Advancing COVID-19 diagnosis with privacy-preserving collaboration in artificial intelligence. *ArXiv.* (2021) 21:698.

105. Alali Y, Harrou F, Sun Y. A proficient approach to forecast COVID-19 spread via optimized dynamic machine learning models. *Scientific Rep.* (2022) 65:798. doi: 10.1038/s41598-022-06218-3

106. Awan MJ, Bilal MH, Yasin A, Nobanee H, Khan NS, Zain AM. Detection of COVID-19 in chest x-ray images: a big data enabled deep learning approach. *Int J Environmen Res Public Health.* (2021) 18:254. doi: 10.3390/ijerph181910147

107. Agarwal A, Henkel R, Huang C-C, Lee M-S. Automation of human semen analysis using a novel artificial intelligence optical microscopic technology. *Andrologia.* (2019) 51:e13440. doi: 10.1111/and.13440

108. Ahmed I, Jeon G. Enabling artificial intelligence for genome sequence analysis of COVID-19 and alike viruses. *Interdisciplin Sci Computat Life Sci.* (2021) 21:1–16. doi: 10.1007/s12539-021-00465-0

109. Babel A, Taneja R, Mondello Malvestiti F, Monaco A, Donde S. Artificial intelligence solutions to increase medication adherence in patients with non-communicable diseases. *Front Digit Health.* (2021) 3:669869. doi: 10.3389/fdgth.2021.669869

110. Anand RS, Stey P, Jain S, Biron DR, Bhatt H, Monteiro K, et al. Predicting mortality in diabetic icu patients using machine learning and severity indices. *AMIA Joint Summits Translation Sci Proceed. AMIA Joint Summ Translat Sci.* (2018) 17:310–319.

111. Baron RJ. Using artificial intelligence to make use of electronic health records less painful-fighting fire with fire. *JAMA Netw Open.* (2021) 4:e2118298. doi: 10.1001/jamanetworkopen.2021.18298

112. Baghel N, Dutta MK, Burget R. Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. *Comput Methods Programs Biomed.* (2020) 197:105750. doi: 10.1016/j.cmpb.2020.105750

113. Bielza C, Larrañaga P, Okamoto H, Brain R. Bayesian networks in neuroscience: a survey. *Jama.* (2014) 8:1–23. doi: 10.3389/fncom.2014.00131

114. Aslam N. Explainable artificial intelligence approach for the early prediction of ventilator support and mortality in COVID-19 Patients. *Computation.* (2022) 22:568. doi: 10.3390/computation10030036

115. Ahn JC, Attia ZI, Rattan P, Mullan AF, Buryska S, Allen AM, et al. Development of the ai-cirrhosis-ecg score: an electrocardiogram-based deep learning model in cirrhosis. *Am J Gastroenterol.* (2022) 117:424–432. doi: 10.14309/ajg.0000000000001617

116. West E. *Intelligence – Based Publications in Radiology From 2000 to 2018.* (2019), pp. 1–3.

117. Alami H, Rivard L, Lehoux P, Hoffman SJ, Cadeddu SBM, Savoldelli M, et al. Artificial intelligence in health care: laying the Foundation for Responsible, sustainable, and inclusive innovation in low- and middle-income countries. *Globalizat Health.* (2020) 16:52. doi: 10.1186/s12992-020-0584-1

118. Schober P, Vetter TR. Linear Regression in Medical Research. *ANESTHESIA & ANALGESIA Statistical Min.* (2021) 132:2020–2021. doi: 10.1213/ANE.0000000000005206

119. Trunfio TA, Scala A, Giglio C, Rossi G, Borrelli A, Romano M, et al. Multiple regression model to analyze the total LOS for patients undergoing laparoscopic appendectomy. *BMC Medical Informat Decision Mak.* (2022) 22:141. doi: 10.1186/s12911-022-01884-9

120. Bonnini S, Borghesi M. *Relationship Between Mental Health and Socio-Economic, Demographic and Environmental Factors in the COVID-19 Lockdown Period — A Multivariate Regression Analysis* (2022).

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 17 August 2023
DOI 10.3389/fpubh.2023.1147709

OPEN ACCESS

EDITED BY
Yingying Xu,
Beihang University, China

REVIEWED BY
Keerti Singh,
The University of the West Indies, Barbados
Zuzana Hajduová,
University of Economics in Bratislava, Slovakia
Baogui Xin,
Shandong University of Science and Technology, China

*CORRESPONDENCE
Syafrawati Syafrawati
✉ syafrawati@ph.unand.ac.id

RECEIVED 19 January 2023
ACCEPTED 31 July 2023
PUBLISHED 17 August 2023

CITATION
Syafrawati S, Machmud R, Aljunid SM and Semiarty R (2023) Incidence of moral hazards among health care providers in the implementation of social health insurance toward universal health coverage: evidence from rural province hospitals in Indonesia. *Front. Public Health* 11:1147709. doi: 10.3389/fpubh.2023.1147709

Incidence of moral hazards among health care providers in the implementation of social health insurance toward universal health coverage: evidence from rural province hospitals in Indonesia

Syafrawati Syafrawati[1]*, Rizanda Machmud[1], Syed Mohamed Aljunid[2,3] and Rima Semiarty[1]

[1]Faculty of Medicine, Andalas University, Padang, Indonesia, [2]Department of Community Medicine, School of Medicine, International Medical University, Kuala Lumpur, Malaysia, [3]International Center for Casemix and Clinical Coding, Faculty of Medicine, National University of Malaysia, Cheras, Malaysia

Objective: To identify the incidence of moral hazards among health care providers and its determinant factors in the implementation of national health insurance in Indonesia.

Methods: Data were derived from 360 inpatient medical records from six types C public and private hospitals in an Indonesian rural province. These data were accumulated from inpatient medical records from four major disciplines: medicine, surgery, obstetrics and gynecology, and pediatrics. The dependent variable was provider moral hazards, which included indicators of up-coding, readmission, and unnecessary admission. The independent variables are Physicians' characteristics (age, gender, and specialization), coders' characteristics (age, gender, education level, number of training, and length of service), and patients' characteristics (age, birth weight, length of stay, the discharge status, and the severity of patient's illness). We use logistic regression to investigate the determinants of moral hazard.

Results: We found that the incidences of possible unnecessary admissions, up-coding, and readmissions were 17.8%, 11.9%, and 2.8%, respectively. Senior physicians, medical specialists, coders with shorter lengths of service, and patients with longer lengths of stay had a significant relationship with the incidence of moral hazard.

Conclusion: Unnecessary admission is the most common form of a provider's moral hazard. The characteristics of physicians and coders significantly contribute to the incidence of moral hazard. Hospitals should implement reward and punishment systems for doctors and coders in order to control moral hazards among the providers.

KEYWORDS

moral hazards, up-coding, readmissions, unnecessary admissions, fraud, physicians, coders, patients

1. Introduction

A moral hazard refers to the possibility of consumers or health care providers abusing a system in order to maximize profits at the expense of other consumers, providers, or the financing community as a whole (1). A moral hazard occurs, for example, when an insured person spends an extra day in the hospital or pays for a procedure that would not have been purchased otherwise (2). In insurance industry, the phenomenon of moral hazard umbrella may be considered as fraud. Insurance fraud would not be possible without asymmetric information—and cheating on insurance companies is deemed immoral—it is referred to as a moral hazard (3). Health insurance fraud can be committed by medical providers, policyholders, or health insurers. Although anyone in the system is capable of committing fraud, healthcare providers are more likely than patients to do so (4, 5).

Some of the healthcare fraud schemes that are frequently discussed in the literature and used to develop fraud detection algorithms or analytics within regulatory entities are as follows: Diagnostic-Related Groups (DRG) creep, unbundling and fragmentation of procedures, up-coding of services, phantom billing, providing excessive services that are not required, kickback schemes, billing for mutually exclusive procedures, duplicate claims and intentional billing errors (6). A number of studies in the world has proven that, provider moral hazard among providers did exist in hospital services (7).

Moral hazard has preoccupied health economics and U.S. health policy for half a century (8). When Medicare providers' payment patterns changed to a prospective Diagnostic Related Groups (DRG) system in the United States, hospitals raised the patient's disease code to a higher level (up-coding). It is aimed at getting the hospital's finance higher than they should be. In private hospitals, the response was stronger. Up-coding or code creeps also occurs in independent medical practices where there is an increase in claim payments, 2.2% from what it should be in 1 year. Hospitals respond to changes in payment patterns by changing the intensity of service provided to patients, severity levels, and market share (7, 9, 10).

Alonazi (11) conducted an audit of the Saudi healthcare system and found the official documents contain the details of various moral hazard measures. Berta et al. (12) examined several types of deviant in Italian hospitals and linked them to hospital efficiency. Deviations in question include up-coding, cream skimming, and readmissions. Debpuur et al. (13) found that the form of moral hazard in the Northern Ghana National Health Insurance is diagnosing simple malaria with complicated malaria, exaggerating the provision of drugs and health services to patients, asking for payments for services that are not provided, and increasing the number of patients receiving health services.

World Health Organization (WHO) estimates the annual global health care expenditure is US$ 5.7 trillion (2008). Each year, 7.29% of that, or an estimated US$ 415 billion, is lost to fraud and errors. South Africa's healthcare system is defrauded between 4 million and 8 billion US Dollars annually. In the UK in 2008–2009, about 3% of National Health Services (NHS) fees were lost to fraud (14). The Centers for Medicare and Medicaid Services (CMS) spent $1.1 trillion on health coverage for 145 million Americans in 2016, $95 billion of which was improper payments related to abuse or fraud (15).

According to a 2009 study, 19.6% of 11.8 million Medicare beneficiaries who were hospitalized from 2003 to 2004 were readmitted within the first month of their hospitalization, costing an estimated $41 billion per year (16). According to Geruso and Layton (2020), upcoding could have cost Medicare $10.5 billion in 2014, or $640 per Medical Advantage enrollee (17). Between July 2009 and June 2010, 139 patients were admitted and treated for preterm labor at a level III center, but none of them delivered preterm. Total hospital charges for the management of these patients were $1,018 589. Unnecessary admissions and treatments for threatened preterm labor are part of clinical practice and contribute to exploding healthcare costs (18).

Indonesia is the world's largest country that aims to achieve Universal Health Coverage through the National Health Insurance. However, the health financing fund was claimed to contribute to budget deficit from USD 200,000,000 in 2014 to USD 450,000,000 in 2016. The moral hazard of health providers has been blamed as one cause of the deficit. In 2015, there were around 175 thousand claims from health services managed by National Health Insurance Administration Agency or recognized by name *BPJS Kesehatan* with a value of 27 million dollars that was detected as fraud, and up to now there have been 1 million claims detected. Nevertheless, so far no independent study was done to assess the real incidence and cause of moral hazards in the National Health Insurance of Indonesia (7, 19, 20). This research can contribute to improve the implementation of Indonesian National Health Insurance by providing scientific evidence on the existence and sources of moral hazards among providers. This will allow the relevant parties to forecast the events and take preventive actions in the future.

On the other hand, systematic review conducted by Pongpirul and Robinson (21) stated that the actors of moral hazard in hospitals could be classified into three categories; those are hospital management, clinicians, and coders (21). However, no study has shown the relationship of moral hazard with the type of hospital, physician, coder, and patient characteristics. The hypothesis we seek to test is that there is a relationship between the characteristics of physicians (age, gender, and specialization), coders (age, gender, education level, number of certificates, and length of service), and patients (age, birth weight, LOS, discharge status, and the severity of the illness) with the incidence of moral hazards. This study aims were to identify the incidence of moral hazards and determinant factors such as physicians, coders and patients characteristics in the implementation of national health insurance in Indonesia.

2. Methods

2.1. Design

We conducted a cross-sectional study on representative Class C hospitals to undertake medical record analysis in West Sumatera Province, Indonesia.

2.1.1. Population

In this study, the population consisted of medical records from inpatients in class C hospitals in West Sumatra. According to data from the Ministry of Health, class C hospitals are the most common type of hospital in West Sumatra. According to data from the Health District Office West Sumatera Province, there are 38 Class C hospitals, 15 of which are government-owned and 23 of which are private (22). A Class C hospital is one that offers four basic medical specializations: surgery, obstetrics and gynecology (OBGYN), pediatrics, and internal medicine.

2.1.2. Sample

A cluster random sampling technique was used to choose the hospitals. Cluster random sampling divides the population into clusters/classes, with the assumption that each class/cluster already has the trait/variation under study. In this study, 6 (six) Class C hospitals were selected, consisting of three government hospitals and three private hospitals. In accordance with the sample size calculation, the minimum sample size for this study is 360 medical records.

Six hospitals receive an equal quantity of samples. In each hospital 60 medical records were selected. The 60 patient medical records will be split into four primary groups of INA-CBG internal disease cases: surgical cases (Group 1), medical cases (Group 4), delivery cases (Group 6), and neonatal cases (Group 8). The number of medical records obtained for each case was 60/4, i.e., 15 medical records per case. The sampling method in this study is shown in Figure 1.

The sample inclusion criteria include medical records which have complete data on 14 casemix variables, including: Patient Data (1). Identification: Patient name, Medical Record Number; (2). Age in years; (3). Age in days; (4). Gender; (5). Date of Birth; (6). Birth Weight (for neonates), Admission Data (7). Date of hospital admission (8). Discharge Date (9). Length of stay (LOS) (10). Discharge Disposition Clinical Data (11). Primary Diagnosis; (12). Secondary Diagnosis; (13). Primary Procedure; (14). Secondary Procedure. Sample exclusions criteria are medical records that are not found, damaged, or cannot be read by an independent coder. Data collection was carried out in January–June 2018.

2.2. Data collection

The data were collected by independent reviewers, namely several senior medical record professionals who did not work in the selected hospitals and have a minimum of 5 to 10 years of experience as coders. Furthermore, the qualification for selecting independent reviewers is that they have attended INA-CBG coding training on a national scale five times or more. The reviewer's job is to go over the medical records of the patient who were chosen as samples. The function of an independent reviewer is to code the patient's illness based on the information in the medical records. Because the independent coder is not involved in the service and management processes at the hospital under review, we consider the results of this coding to be the gold standard for medical coding. Furthermore, the reviewers gathered secondary data in the form of the characteristics of the coder, clinician, and patient, which were the study's independent factors.

2.3. Study outcomes

2.3.1. Moral hazard

Three indicators of moral hazards are used in this study: up-coding, readmission and unnecessary admission. These variables were derived from systematic review and a pilot study to identify the main moral hazard indicators in hospitals.

Up-coding is the mismatch between the diagnosis code and procedure written in the medical records which causes an increase in hospital reimbursement (7, 23–25). In this study, an Independent Senior Coder (ISC) reviewed each of the medical records. The codes were entered into INA-CBGs software to determine the hospital tariff. Furthermore, researchers gathered data on the results of medical coding executed by the hospital coder (original codes) and their tariff. The tariff based on the original codes was compared with the codes from ISC. If the tariff from the hospital coder's work is higher than ISC codes, then the case is considered as up-coding.

Readmission is an event of patient service where the same discharged inpatient is brought back for hospitalization to undergo the same disease treatment after a period of <30 days (26). In this study, we reviewed the medical records and if the is hospitalized in the same hospital for the same disease they had previously been treated and discharged <30 days, then it is classified as a readmission case.

Unnecessary admission is a hospitalization case where there is no significant reason for the patient to be treated when they were first admitted to the hospital (26). In this study, unnecessary admission is defined as any admission with a length of stay (LOS) of 2 days and below, and the patient is discharged well (not dead). However, admissions that ended in death are not considered unnecessary admissions.

2.3.2. Characteristics of patients

Patient data is an important factor to determine INA-CBGs tariff. Patient data consist of demographic data, admission data, and clinical data. So far, there were no studies that look at the relationship between patient variables with moral hazard. Patient variables in this study were patients' age, birth weight, LOS, discharge status, and the severity of the patient's illness.

2.3.3. Characteristics of coders

A coder is a person who assigned the diagnoses and procedure codes and enters the minimum data set into INA-CBGs software

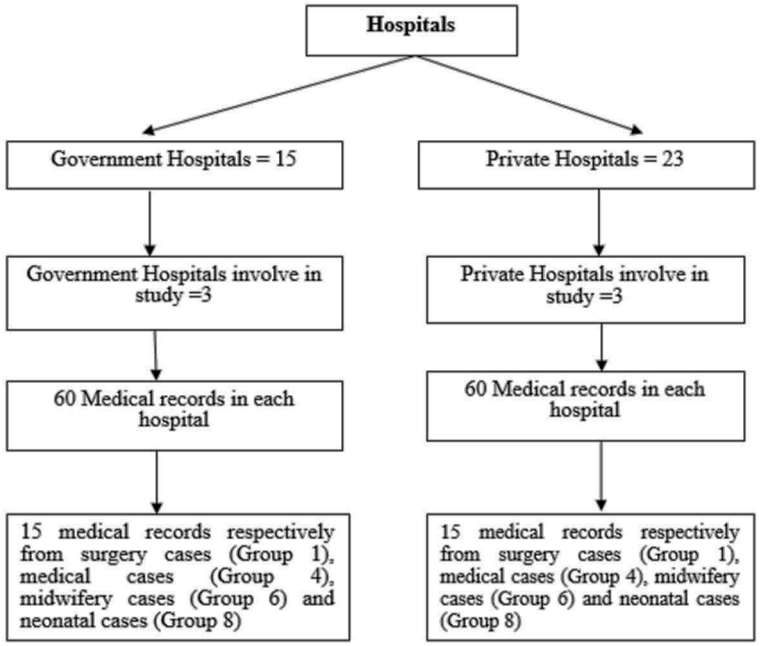


FIGURE 1
Sample selection process.

in order to produce INA-CBGs tariff. The coder's qualification is very decisive for coding quality. Coders' characteristics in this study were age, gender, education level, number of certificates, and length of service.

2.3.4. Characteristics of physicians

In terms of health care provided in hospitals, a physician is a person who has responsibility for patient care. The physician also has the potential to perform moral hazard by increasing admission volume, changing the intensity of care, and exaggerating (21). Physician variables in this study were age, gender, and specialization (medical, surgical, OBGYN, pediatric).

2.4. Statistical analysis

The incidence of moral hazard and characteristics of doctors, coders, and patients are described in the frequency distribution table. Incident moral hazard consists of up-coding, readmission, and unnecessary admission. The physician's characteristics include the physician's age, physician's gender, and physician's specialization (medical, surgical, OBGYN, and pediatric). The characteristics of the coder consist of the coder's age, coder's gender, education level of coders, the number of coder's certificates, and coder's length of service. Patient characteristics consist of patients' age, birth weight, LOS, discharge status, and the severity of the patient's illness.

We used multilevel logistic regression analysis to examine the contributions of characteristics of the patient, coder, and physician to the incidence of moral hazards. The following multilevel model was used (27).

$$\frac{P\left(y=1\middle|x_{ij},\eta_{0j}\right)}{P\left(y=0\middle|x_{ij},\eta_{0j}\right)}=\beta_0+\beta_1+\ldots\beta_{\mathrm{k}}X_{\mathrm{kij}}+\eta_{0\mathrm{j}}+e_{\mathrm{ij}}$$

In the presence of more than one explanatory variable, logistic regression is used to calculate the odds ratio. With the exception that the response variable is binomial, the approach is quite similar to multiple linear regression. The impact of each variable on the odds ratio of the observed event of interest is the result (28).

2.5. Ethics approval

Ethical approval for this study was obtained from Faculty of Medicine Andalas University (No. 052/KEP/FK/2018).

3. Results

3.1. Incidence of moral hazard

Detailed indicators of moral hazard are presented in Table 1. The most common type of moral hazard was possible unnecessary admission (17.8%), followed by up-coding (11.9%) and readmission (2.8%).

Unnecessary admissions were up to 4.2% more common among neonates group. Meanwhile, deliveries group dominated up coding cases by as much as 2.8%. Readmissions were more common in neonatal groups and female reproductive system groups. The full results can be seen in Table 2.

TABLE 1 Incidence of moral hazard.

Indicators of moral hazard		Numbers	%
Up-coding	Yes	43	11.9
	No	317	88.1
Readmission	Yes	10	2.8
	No	250	97.2
Unnecessary admission	Yes	64	17.8
	No	296	82.8
Total		360	100

3.2. The characteristic of physicians, coders, and patients

Table 3 illustrates the characteristics of physicians, coders and patients. The average age of physicians was 41 years old. It means that the physicians involved in this study were mostly young. Similarly the average age of coders was 31.32 years old. Meanwhile, the average age of patients was 26, 49 years old, and average LOS was 4 days.

Most of physicians' were male and one third of them were specialized in OBGYN (31.4%). Half (50.8%) of coders have had <4 years work experience and more than half coder (67.5%) have at less only one INA CBG coding training. Most of patients (71.4%) are females, with the discharge status dominated discharge home (93.6%). Most of the infant patients have not experienced low birth weight (84.6%) as illustrated in Table 4.

It could be seen from multivariate analysis that physician's age, physicians' specialization, coders' length of service, and LOS have significant relationship with incidence of moral hazard. The most remarkable influence on moral hazard cases is physicians specialization variable. To put it simply older physicians, medical specialization, coders with less length of service, and long LOS had a significant relationship with the incidence of moral hazard. The full results can be seen in Table 5.

4. Discussion

Indonesia offers significant funding for JKN implementation. According to the *BPJS Kesehatan* financial report, health insurance expenses totalled 6.364 billion US dollar in 2018. Several rules to prevent moral hazard or fraud have also been implemented, such as the release of Regulation of the Minister of Health no. 36 of 2015 on hospital fraud prevention. However, no research has been conducted to demonstrate the efficiency of these preventative measures against moral hazard situations in hospitals.

This study found the incidence of unnecessary admission was the highest moral hazard indicator at 17.8%. This finding is higher than in other studies elsewhere. According to Mosadeghrad and Isfahani (30) research on the measurement of unnecessary patient admissions in Iranian hospitals, 2.7% of hospital admissions were considered unacceptable and unnecessary. The highest unnecessary patients' admissions in hospital were 11.8%, and the lowest were 0.3%.

Unnecessary admission is "an admission that provides no significant benefit to the patient or provides a benefit that could have been obtained at a lower level of care (31). In this study, the term "unnecessary admission" refers to patients who are hospitalized for 1 to 2 days with a non-dead discharge status. Unnecessary admissions mostly occurred in the Neonatal Group and Deliveries Group. Other studies elsewhere on unnecessary admissions were mostly found in the emergency departments (32–34).

A variety of patient-related factors (e.g., age, disease severity, method of payment, and route and time of admission), physicians, and the hospital and its diagnostic facilities and technology influence the unnecessary admission of patients to the hospital. Unnecessary hospitalization increases nosocomial infections, morbidity, and mortality, and reduces patient satisfaction and hospital productivity (35–39).

Previous researchers proposed several strategies for reducing avoidable hospital admissions, including expanding the primary health care network, reducing hospital beds, implementing an effective and efficient patient referral system, using a fixed provider payment method, promoting residential and social services care at the macro level, establishing a utilization management committee, using the appropriateness evaluation protocol, establishing short-stay units, and establishing a patient referral system (30, 33, 40, 41).

Indonesia has implemented several of these strategies in its health care system, such as implementing a patient referral system. The National Health Insurance Administration Agency has a tiered referral system that must be implemented by health insurance participants, social health insurance companies, and health facility providers. This tiered referral system operates on a hierarchical basis, beginning with primary health facilities (the closest to the community) and progressing to secondary and tertiary health facilities. Referral to second-level health facilities can only be administered by a first-level health facility (42).

With the existence of a referral system, where there are criteria such as health services in primary health care facilities that can be referred directly to tertiary health care facilities only for cases that have been diagnosed and a treatment plan has been established, a repeat service and is only available in tertiary health care facilities, reducing the incidence of unnecessary admissions because there is a system that must be followed, unnecessary admissions will be reduced. This system is strengthened by the existence of a policy that states that if a health facility does not adopt a referral system, *BPJS Kesehatan* will conduct re-credentialing on the health facility's performance, which may have an impact on future collaboration.

However, hospitals must strengthen management to avoid unnecessary admissions by establishing a utilization management committee and implementing the appropriate evaluation protocol.

The second type of moral hazard found in this study was up-coding (11.9%). In Germany, up-coding occurs at 1% of inpatients' payments (29). Another study found a fairly high incidence of up-coding, estimating that 18.5% annual reimbursed claims for Present on Admission (POA) infections were up-coded hospital-acquired infections (HAIs) (43).

TABLE 2 Percentage of moral hazard types based on casemix main group.

Moral hazard	Casemix main groups	Total	%	Explanation
Up-coding (43)	O	10	2.8	Deliveries group
	W	6	1.7	Female reproductive system groups
	P	4	1.1	Newborns and neonates groups
	K	4	1.1	Digestive system group
	G	3	0.8	Central nervous system groups
	I	3	0.8	Cardiovascular system groups
	L	3	0.8	Skin, subcutaneous tissue and breast group
	M	3	0.8	Musculoskeletal system and connective tissue groups
	D	2	0.5	Haemopoeitic and immune system groups
	C	1	0.3	Myeloproliferative system and neoplasms groups
	E	1	0.3	Endocrine system, nutrition and metabolism groups
	J	1	0.3	Respiratory system groups
	N	1	0.3	Nephro-urinary system groups
	U	1	0.3	Ear, nose, mouth and throat groups
Readmission (10)	P	2	0.5	Newborns and neonates group
	W	2	0.5	Female reproductive system groups
	G	1	0.3	Central nervous system groups
	I	1	0.3	Cardiovascular system group
	J	1	0.3	Respiratory system groups
	K	1	0.3	Digestive system group
	L	1	0.3	Skin, subcutaneous tissue and breast groups
	U	1	0.3	Ear, nose, mouth and throat groups
Possible unnecessary admission (64)	P	15	4.2	Newborns and neonates group
	O	14	3.9	Deliveries group
	U	10	2.8	Ear, nose, mouth and throat groups
	M	6	1.7	Musculoskeletal system and connective tissue groups
	G	4	1.1	Central nervous system groups
	H	3	0.8	Eye and adnexa groups
	L	3	0.8	Skin, subcutaneous tissue and breast groups
	W	3	0.8	Female reproductive system groups
	D	2	0.5	Haemopoeitic and immune system groups
	K	2	0.5	Digestive system group
	I	1	0.3	Cardiovascular system groups
	N	1	0.3	Nephro-urinary system groups

Hospitals in Germany have up-coded at least 12,000 premature babies and received additional reimbursements totalling more than 100 million Euros since the implementation of DRG. Currently, approximately 2,000 up-coding generate an additional 20 million Euros per year (44).

Up-coding is the practice of classifying a patient in a DRG that results in a higher reimbursement or shifting a patient's DRG to another DRG that results in a higher payment from the third-party provider (25, 45). There are two primary methods for detecting potential DRG up-coding: (1) auditing by recoding the original medical charts, and (2) comparing historical claim data to detect an increase in the percentage of higher-cost DRGs (24).

Previous studies have indicated that DRG up-coding by private providers can be intentional (46). A code audit is the most reliable method of detecting DRG up-coding. Experienced health-information managers recode the original medical chart and then compare the new codes to the codes originally submitted by the hospital in code audit (46). Other research indicates that audits with fines can reduce up-coding while not necessarily inducing more honesty (47).

10.3389/fpubh.2023.1147709

TABLE 3 Characteristics of physicians, coders, and patients.

Variables	Mean (±SD)	Median	Max	Min
Physicians; age	41.04 (±9,037)	39	72	31
Coders; age	31.32 (±5,244)	30	40	26
Patients' age	26.49 (±22,387)	25.5	86	0
LOS	4.18 (±2,213)	4	20	1

Another qualitative study on up-coding discovered that the Deliveries Group (2.8%) had the highest percentage of up-coding, followed by Female Reproductive System Groups (1.7%). Upcoding could result in a loss of IDR 154,626,000, or 9% of hospital revenue (48).

That study also discovered that the reasons for up-coding can be divided into three categories: (1) hospital-related; this occurred due to a lack of defined coding criteria. The hospital also did not know the flow of coordination between the teams constituted to tackle the problem of coding conflicts between the hospital and *BPJS Kesehatan*. (2) related to doctors; and (3) related to coders. Doctors frequently did not understand the coding standards. From the doctor's perspective, the disease's symptoms could also be incorporated into medical coding, but they couldn't. Furthermore, the coders occasionally have problems reading, and the doctor's handwriting and untranslated abbreviations are illegible (48).

To avoid human error in up-coding, doctors and coders should receive medical coding training to reduce doctor misspecifications or coder misunderstandings.

So far, the government's efforts in reducing moral hazard, including up-coding, have included the signing of an agreement or memorandum of understanding (MoU) between the *BPJS Kesehatan* and the Director of the Hospital, which includes the "Declaration of Absolute Responsibility Submission of Health Services Claims" and the "Statement of Claims by the Team Hospital Fraud Prevention," which specifies that the hospital director is accountable for submitting claims files that are devoid of fraud or moral hazard. If there is fraud, the director is willing to face legal consequences.

This study discovered a low number of readmission incidents, with only 2.8%. Readmission cases were confirmed more in group P (Newborns and Neonates Group) and group W (Female reproductive system Groups) with two cases each.

Hospital readmissions can be defined as admissions to hospitals or other health care facilities arranged within a specific period of time following a hospital stay. Readmissions can also be defined as returning to the hospital within 30 days of being discharged (at first time), allowing the hospital to receive multiple reimbursements for the same treatment (12, 49, 50). The following factors contribute to readmissions: Inability to recognize the seriousness of the patient's illness, inability to appropriately address the patient's illness, patients being discharged from the hospital prematurely, and a lack of control over the hospital (12, 51, 52).

Furthermore, in research conducted by Auger et al. (53) on medical record review, 15% of readmissions were classified as unplanned and preventable. Researchers and policymakers concluded that a significant proportion of readmissions were

TABLE 4 Characteristics of physicians, coders, and patients (categorical data).

Variabel	F	%
Physician		
Physician's sex		
Male	244	**67.8**
Female	116	32.2
Physician's specialization		
Surgery	78	21.7
Medical (Internal Medicine, Ophthalmology, Cardiology, ENT (Ear, Nose and Throat), Pulmonology	67	18.6
Obstetric and genecology	113	**31.4**
Pediatric	102	28.3
Coder		
Coder's Sex		
Male	0	0
Female	360	**100**
Coder education		
Lower than diploma	0	0
Diploma and higher	360	**100**
Number of coder's training		
One and none	243	**67.5**
More than one	117	32.5
Coder length of services		
Less than 4 years	183	**50.8**
More than 4 years	177	49.2
Patients		
Patient's sex		
Male	103	28.6
Female	257	**71.4**
Birth weight (neonates)		
Low birth weight	14	15.4
Normal	77	**84.6**
Discharge status		
Discharge home	337	**93.6**
Transferred to other hospitals	11	3.1
Discharge against medical advice	6	1.7
Death	6	1.7
Primary diagnosis		
A00-B99	8	2.2
C00-D48	25	6.9
D50-D89	6	1.7
E00-E90	6	1.7
H00-H59	2	0.6

(Continued)

TABLE 4 (Continued)

Variabel	F	%
I00-I99	23	6.4
J00-J99	22	6.1
K00-K93	32	8.9
L00-L99	5	1.4
M00-M99	6	1.7
N00-N99	10	2.8
O00-O99	105	**29.2**
P00-P96	89	**24.7**
Q00-Q99	2	0.6
R00-R99	6	1.7
S00-T98	8	2.2
Z00-Z99	5	1.4
Primary procedure		
No procedure	139	**38.6**
01–05	1	0.3
06–07	1	0.3
08–16	3	0.8
21–29	17	4.7
30–34	2	0.6
35–39	2	0.6
40–41	1	0.3
42–54	21	5.8
55–59	1	0.3
65–71	8	2.2
72–75	89	**24.7**
76–84	19	5.3
85–86	21	5.8
87–99	35	9.7
Severity level of illness		
Severity level 3	26	7.2
Severity level 2	57	15.8
Severity level 1	277	**76.9**
Total	**360**	**100**

The bold values indicate the highest percentage of each variable.

caused by healthcare system failures—whether due to inadequate treatment during the initial hospitalization or a failure of care coordination after hospital discharge. Therefore, it is necessary to have policies to reduce inappropriate readmissions because hospitals receive additional payments when patients are readmitted (54).

Several interventions have succeeded in reducing readmission rates for discharged patients. These interventions include: patient needs assessment, medication reconciliation, patient education, timely outpatient appointments, and telephone follow-up. The impact of the intervention on the readmission rate is proportional to the number of components performed. This means that interventions with single component treatments are unlikely to reduce readmissions significantly (55).

Another finding in our study was physicians' age and specialization, coder's length of service, and LOS was the determinant factor of moral hazard in hospitals.

Older physicians are 1.037 times more likely than younger physicians to be associated with moral hazards (POR = 1.037). Other studies found that male physicians and older physicians were more likely to commit fraud, waste, and abuse on Medicare (56, 57). Based on the in-depth research conducted, it is stated, before beginning inpatient care, younger physicians learned the fundamentals of rules. As a result, their knowledge of coding rules keeps them more aware of moral hazards than older physicians. Older physicians may be resistant to new patient treatment rules, particularly coding rules that they believe are unfair to them. It is suggested that every CME (Continuing Medical Education) unit in the Faculty of Medicine should include Moral Hazard material in its activities in order to increase knowledge and skills, as well as develop doctors' attitudes so that they can always carry out their profession properly and correctly, and to help physicians understand the consequences of moral hazards and avoid them in the interest of their patient's health.

We also discovered that the specialization of physicians can influence the occurrence of moral hazards. Based on the study's findings, medical specializations are 2.373 times more likely to perform moral hazards rather than surgical, pediatrics, and OBGYN specializations (POR = 2.373). Previous research has found that the specialization of physicians influences the occurrence of fraud. According to Chen (57), physicians in certain specialties (such as family medicine, psychiatry, internal medicine, anesthesiology, surgery, and OBGYN) are more likely to commit Medicare fraud, waste, and abuse.

Pediatric specialization is used as the standard for calculating baselines in this study because the algorithm for compiling the INA-CBGs code in cases of pediatric is more complex, making manipulation difficult. According to this study, medical specialization is more likely to cause moral hazard than other specializations. Previous studies discovered Family medicine physicians and psychiatrists departed are more likely to commit fraud. This is because fraud is easier to commit when the risk of malpractice suits is very low, such as in the fields of family medicine and psychiatry. The study also explains that surgeons have the highest proportion of doctors who face malpractice claims based on their specialization (58).

BPJS Kesehatan is expected to be more stringent in inspecting cases of moral hazard to medical specialists and to continue to educate and raise awareness among physicians about the potential moral hazard in their medical practices.

We discovered that the length of service of the coder was a determinant of moral hazards. A beginner coder was 2.237 times more likely than an experienced coder to commit moral hazard (POR = 2.237). The length of service corresponds to the opportunity to receive training. Because the INA-CBG coding rules are not taught in detail in their studies, beginner coders have limitations in understanding the hospital coding regulations. Hospitals should provide regular training, particularly for new coders.

TABLE 5 Multiple logistic regression analysis of factors influencing moral hazards.

Variabel	B	S.E.	Wald	*p*-value	POR	95.0% C.I. for EXP(B)
Physician's age	0.036	0.018	4.054	0.044	1.037	1,001–1,073
Physician's specialization						
Surgery	0.752	0.356	4.459	0.035	2.121	1.056–4.262
Medical	0.864	0.377	5.257	0.022	2.373	1.134–4.262
Obstetric and gynecology	0.219	0.338	0.421	0.517	1.245	0,642–2,415
Pediatric					1	
Coder length of services	0.805	0.255	9.942	0.002	2.237	1,356–3,691
LOS	0.222	0.069	10.397	0.001	1.249	1,091–1,430

The LOS is the final factor discovered in this study that influences moral hazard in hospitals. Moral hazards are 1.249 times more likely to occur in long lengths of stay than short lengths of stay (POR = 1.249). Patients who require more treatment spend more days in hospitals. More hospital resources will be deployed as a result. Some hospitals are willing to take moral hazard, such as up-coding if they believe the INA-CBG tariff is insufficient for patient care. We recommend that the *BPJS Kesehatan* conduct more audits of hospitals with higher LOS.

The study's findings are highly encouraging, as it is well-known that senior physicians, medical specialists, coders with shorter lengths of service, and patients with longer lengths of stay. This discovery should encourage hospitals and insurance companies to be more cautious and pay more attention to audits of patient medical records containing these variable factors.

5. Strengths

In this study, each medical record was examined by an Independent Senior Coder (ISC). ISC comes from a higher-level hospital than the one being analyzed, where the coders are accustomed to coding more difficult and complex cases, and they have had national-level training more than five times.

6. Limitations

The operational definition of unnecessary admission in this study is a hospitalization of <2 days, and discharges, which are not due to death. No doctor conducts a thorough examination of the unnecessary admission case. As a result, some cases of unnecessary admission may be considered necessary admission because they are not further evaluated by medical experts. The next study is expected to include a physician reviewing each case to determine whether hospitalization or admission is required.

7. Conclusion

This study revealed that the most common moral hazard is unnecessary admission, followed by up-coding and readmission. The factors significantly associated with moral hazard are physicians' age, physicians' specialization, coders' length of service, and LOS. The main factor that most has a role in moral hazard is the physician's specialization. It is suggested to the hospitals conduct training for physicians and coders about coding rules in casemix system in the hospital.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

SS, RM, SA, and RS contributed to conception and design of the study. SS organized the database and wrote the first draft of the manuscript. SS, RM, and SA performed the statistical analysis and wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by Ministry of Health of Indonesia for providing the financial support for this study (HK.03.01/I/987/2019).

Acknowledgments

We acknowledge Ministry of Health of Indonesia for providing the financial support for this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Andargie G. *Introduction to Health Economics*. Gondar: University of Gondar (2008).

2. Nyman JA. Is "moral hazard" inefficient? The policy implications of a new theory. *Health Affairs.* (2004) 23:194–9. doi: 10.1377/hlthaff.23.5.194

3. Van Wolfen J, Inbar Y, Zeelenberg M. *Moral Hazard in the Insurance Industry*. Tilburg: Tilburg University (2013).

4. Peck S, McKenna L. Fraud in healthcare: a worldwide concern. *Health Manag.* (2017) 17:124–6. Available online at: https://healthmanagement.org/c/healthmanagement/issuearticle/fraud-in-healthcare-a-worldwide-concern

5. Villegas-Ortega J, Bellido-Boza L, Mauricio D. Fourteen years of manifestations and factors of health insurance fraud, 2006–2020: a scoping review. *Health Justice.* (2021) 9:1–23. doi: 10.1186/s40352-021-00149-3

6. Kumaraswamy N, Markey MK, Ekin T, Barner JC, Rascati K. Healthcare fraud data mining methods: a look back and look ahead. *Perspect Health Inform Manag.* (2022) 19. Available online at: https://perspectives.ahima.org/page/healthcare-fraud-data-mining-methods-a-look-back-and-look-ahead

7. Dafny LS. How Do Hospitals respond to price changes? *Am Econ Rev.* (2005) 95:1525–47. doi: 10.1257/000282805775014236

8. Robertson CT, Yuan A, Zhang W, Joiner K. Distinguishing moral hazard from access for high-cost healthcare under insurance. *PLoS ONE.* (2020) 15:e0231768. doi: 10.1371/journal.pone.0231768

9. Seiber EE. Physician code creep: evidence in medicaid and state employee health insurance billing. *Health Care Financ Rev.* (2007) 28:83.

10. Ellis RP, McGuire TG. Hospital response to prospective payment: moral hazard, selection, and practice-style effects. *J Health Econ.* (1996) 15:257–77. doi: 10.1016/0167-6296(96)00002-1

11. Alonazi WB. Fraud and Abuse in the Saudi healthcare system: a triangulation analysis. *Inquiry J Health Care Organ Prov Finan.* (2020) 57:0046958020954624. doi: 10.1177/0046958020954624

12. Berta P, Callea G, Martini G, Vittadini G. The effects of upcoding, cream skimming and readmissions on the Italian hospitals efficiency: a population-based investigation. *Econ Model.* (2010) 27:812–21. doi: 10.1016/j.econmod.2009.11.001

13. Debpuur C, Dalaba MA, Chatio S, Adjuik M, Akweongo P. An exploration of moral hazard behaviors under the national health insurance scheme in Northern Ghana: a qualitative study. *BMC Health Serv Res.* (2015) 15:1–9. doi: 10.1186/s12913-015-1133-4

14. Jones B, Jing A. Prevention not cure in tackling health-care fraud. *World Health Orga Bullet World Health Org.* (2011) 89:858. doi: 10.2471/BLT.11.021211

15. Drabiak K, Wolfson J. What should health care organizations do to reduce billing fraud and abuse? *AMA Journal of Ethics.* (2020) 22:221–31. doi: 10.1001/amajethics.2020.221

16. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med.* (2009) 360:1418–28. doi: 10.1056/NEJMsa0803563

17. Geruso M, Layton T. Upcoding: evidence from Medicare on squishy risk adjustment. *J Polit Econ.* (2020) 128:984–1026. doi: 10.1086/704756

18. Lucovnik M, Chambliss LR, Garfield RE. Costs of unnecessary admissions and treatments for "threatened preterm labor". *Am J Obst Gynecol.* (2013) 209:217.e1–3. doi: 10.1016/j.ajog.2013.06.046

19. BPJS. *Management Program Report Year 2016 and Financial Statement 2016*. BPJS Kesehatan (2017). Available online at: https://www.bpjs-kesehatan.go.id/#/formasi-public-detail?slug=d882051c-6e27-4cdc-b19c-530e19f1c759 (accessed October 18, 2022).

20. Rahma P. *Strengthening the Role of the JKN Fraud Prevention Team in the Regions to Control Fraud. Pusat Kebijakan dan Manajemen Kesehatan Fakultas Kedokteran.* Yogyakarta*:* Universitas Gadjah Mada (2019).

21. Pongpirul K, Robinson C. Hospital manipulations in the DRG system: a systematic scoping review. *Asian Biomed.* (2013) 7:301–10. doi: 10.5372/1905-7415.0703.180

22. Indonesian Ministry of Health. Hospital Data Online. Jakarta. Hospital system Online. In: *Health IMo*, editor. Jakarta. Indonesia. (2017).

23. Lorence DP, Richards M. Variation in coding influence across the USA. Risk and reward in reimbursement optimization. *J Manag Med.* (2002) 16:422–35. doi: 10.1108/02689230210450981

24. Silverman E, Skinner J. Medicare upcoding and hospital ownership. *J Health Econ.* (2004) 23:369–89. doi: 10.1016/j.jhealeco.2003.09.007

25. Barros P, Braun G. Upcoding in a national health service: the evidence from Portugal. *Health Econ.* (2017) 26:600–18. doi: 10.1002/hec.3335

26. Mills A. Health care systems in low and middle income countries. *N Engl J Med.* (2014) 370:552–7. doi: 10.1056/NEJMra1110897

27. Merlo J, Yang M, Chaix B, Lynch J, Råstam L. A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people. *J Epidemiol Commun Health.* (2005) 59:729–36. doi: 10.1136/jech.2004.023929

28. Sperandei S. Understanding logistic regression analysis. *Biochem Med.* (2014) 24:12–8. doi: 10.11613/BM.2014.003

29. Lüngen M, Lauterbach K. Upcoding-a risk for the use of diagnosis-related groups. *Deutsche Med Wochenschrift (1946).* (2000) 125:852–6. doi: 10.1055/s-2000-7019

30. Mosadeghrad AM, Isfahani P. Unnecessary hospital admissions in Iran: a systematic review and meta-analysis. *Tehran Univ Med J TUMS Pub.* (2019) 77:392–400.

31. Eriksen B, Kristiansen I, Nord E, Pape J, Almdahl S, Hensrud A, et al. The cost of inappropriate admissions: a study of health benefits and resource utilization in a department of internal medicine. *J Intern Med.* (1999) 246:379–87. doi: 10.1046/j.1365-2796.1999.00526.x

32. Ben-Assuli O, Leshno M, Shabtai I. Using electronic medical record systems for admission decisions in emergency departments: examining the crowdedness effect. *J Med Syst.* (2012) 36:3795–803. doi: 10.1007/s10916-012-9852-0

33. Carlill G, Gash E, Hawkins G. Preventing unnecessary hospital admissions: an occupational therapy and social work service in an accident and emergency department. *Br J Occup Therapy.* (2002) 65:440–5. doi: 10.1177/030802260206501002

34. Collins SP, Pang PS, Fonarow GC, Yancy CW, Bonow RO, Gheorghiade M. Is hospital admission for heart failure really necessary? The role of the emergency department and observation unit in preventing hospitalization and rehospitalization. *J Am Coll Cardiol.* (2013) 61:121–6. doi: 10.1016/j.jacc.2012.08.1022

35. O'Cathain A, Knowles E, Turner J, Hirst E, Goodacre S, Nicholl J. Variation in avoidable emergency admissions: multiple case studies of emergency and urgent care systems. *J Health Serv Res Policy.* (2016) 21:5–14. doi: 10.1177/1355819615596543

36. Pope I, Burn H, Ismail SA, Harris T, McCoy D, A. qualitative study exploring the factors influencing admission to hospital from the emergency department. *BMJ Open.* (2017) 7:e011543. doi: 10.1136/bmjopen-2016-011543

37. Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. *Jama.* (1992) 268:2388–94. doi: 10.1001/jama.1992.03490170060026

38. McGregor MJ, Reid RJ, Schulzer M, Fitzgerald JM, Levy AR, Cox MB. Socioeconomic status and hospital utilization among younger adult pneumonia admissions at a Canadian hospital. *BMC Health Serv Res.* (2006) 6:1–10. doi: 10.1186/1472-6963-6-152

39. Pileggi C, Bianco A, Di Stasio S, Angelillo I. Inappropriate hospital use by patients needing urgent medical attention in Italy. *Public Health.* (2004) 118:284–91. doi: 10.1016/j.puhe.2003.06.002

40. Richards III F, Pitluk H, Collier P, Powell S, Dion C, Struchen-Shellhorn W, et al. Reducing unnecessary Medicare hospital admissions for chest pain in Arizona and Florida. *Prof Case Manag.* (2008) 13:74–84. doi: 10.1097/01.PCAMA.0000314177.01661.b3

41. Arab-Zozani M, Pezeshki MZ, Khodayari-Zarnaq R, Janati A. Inappropriate rate of admission and hospitalization in the iranian hospitals: a systematic review and meta-analysis. *Value in Health Regional Issues.* (2020) 21:105–12. doi: 10.1016/j.vhri.2019.07.011

42. Handayani PW, Saladdin IR, Pinem AA, Azzahro F, Hidayanto AN, Ayuningtyas D. Health referral system user acceptance model in Indonesia. *Heliyon.* (2018) 4:e01048. doi: 10.1016/j.heliyon.2018.e01048

43. Bastani H, Goh J, Bayati M. Evidence of upcoding in pay-for-performance programs. *Manage Sci.* (2019) 65:1042–60. doi: 10.1287/mnsc.2017.2996

44. Jürges H, Köberlein J. What explains DRG upcoding in neonatology? The roles of financial incentives and infant health. *J Health Econ.* (2015) 43:13–26. doi: 10.1016/j.jhealeco.2015.06.001

45. Simborg DW. DRG creep: a new hospital-acquired disease. *Mass Medical Soc.* (1981) 3:1602–4. doi: 10.1056/NEJM198106253042611

46. Luo W, Gallagher M, editors. Unsupervised DRG upcoding detection in healthcare databases. *2010 IEEE International Conference on Data Mining Workshops* 2010: IEEE. (2010). doi: 10.1109/ICDMW.2010.108

47. Groß M, Jürges H, Wiesen D. The effects of audits and fines on upcoding in neonatology. *Health Econ.* (2021) 30:1978–86. doi: 10.1002/hec.4272

48. Syafrawati S, Machmud R, Aljunid SM, Semiarty R. Incidence and root cause of upcoding in the implementation of social health insurance in rural province hospital in Indonesia. *Asia Pacific Fraud J.* (2020) 5:56–61. doi: 10.21532/apfjournal.v5i1.135

49. Hosseinzadeh A, Izadi M, Verma A, Precup D, Buckeridge D, editors. Assessing the predictability of hospital readmission using machine learning. *Twenty-FIfth IAAI Conference.* (2013). doi: 10.1609/aaai.v27i2.18995

50. Stone JL, Hoffman G. *Medicare Hospital Readmissions: Issues, Policy Options and PPACA*. Washington, DC: Congressional Research Service (2010).

51. Niu Y. *Regression Models for Readmission Prediction Using Electronic Medical Records*. Detroit, MI: Wayne State University (2013).

52. Khawaja FJ, Shah ND, Lennon RJ, Slusser JP, Alkatib AA, Rihal CS, et al. Factors associated with 30-day readmission rates after percutaneous coronary intervention. *Arch Intern Med.* (2012) 172:112–7. doi: 10.1001/archinternmed.2011.569

53. Auger KA, Ponti-Zins MC, Statile AM, Wesselkamper K, Haberman B, Hanke SP. Performance of pediatric readmission measures. *J Hospital Med.* (2020) 15:723–6. doi: 10.12788/jhm.3521

54. Cram P, Wachter RM, Landon BE. Readmission reduction as a hospital quality measure: time to move on to more pressing concerns? *JAMA*. (2022) 328:1589–90. doi: 10.1001/jama.2022.18305

55. Kripalani S, Theobald CN, Anctil B, Vasilevskis EE. Reducing hospital readmission: current strategies and future directions. *Annu Rev Med.* (2014) 65:471. doi: 10.1146/annurev-med-022613-090415

56. Pande V, Maas W. Physician medicare fraud: characteristics and consequences. *Int J Pharm Healthcare Mark.* (2013) 7:8–33. doi: 10.1108/17506121311315391

57. Chen A, Blumenthal DM, Jena AB. Characteristics of physicians excluded from US Medicare and state public insurance programs for fraud, health crimes, or unlawful prescribing of controlled substances. *JAMA Network Open.* (2018) 1:e185805–e. doi: 10.1001/jamanetworkopen.2018.5805

58. Jena AB, Seabury S, Lakdawalla D, Chandra A. Malpractice risk according to physician specialty. *New Eng J Med.* (2011) 365:629–36. doi: 10.1056/NEJMsa1012370

frontiers | Frontiers in Public Health

TYPE Systematic Review
PUBLISHED 24 November 2023
DOI 10.3389/fpubh.2023.1151504

OPEN ACCESS

EDITED BY
Yingying Xu,
Beihang University, China

REVIEWED BY
Mario Dioguardi,
University of Foggia, Italy
Theo Vos,
University of Washington, United States

*CORRESPONDENCE
Wenjia Chen
✉ wenjiach@nus.edu.sg

RECEIVED 26 January 2023
ACCEPTED 04 October 2023
PUBLISHED 24 November 2023

CITATION
Chen W, Wong NCB, Wang Y, Zemlyanska Y, Butani D, Virabhak S, Matchar DB, Prapinvanich T and Teerawattananon Y (2023) Mapping the value for money of precision medicine: a systematic literature review and meta-analysis.
Front. Public Health 11:1151504.
doi: 10.3389/fpubh.2023.1151504

Mapping the value for money of precision medicine: a systematic literature review and meta-analysis

Wenjia Chen[1]*, Nigel Chong Boon Wong[1], Yi Wang[1], Yaroslava Zemlyanska[1], Dimple Butani[2], Suchin Virabhak[3], David Bruce Matchar[3,4], Thittaya Prapinvanich[5] and Yot Teerawattananon[1,2]

[1]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore, [2]Health Intervention and Technology Assessment Program (HITAP), Ministry of Public Health, Bangkok, Thailand, [3]Precision Health Research, Singapore (PRECISE), Singapore, Singapore, [4]Duke-NUS Medical School, Singapore, Singapore, [5]Yale-NUS College, Singapore, Singapore

Objective: This study aimed to quantify heterogeneity in the value for money of precision medicine (PM) by application types across contexts and conditions and to quantify sources of heterogeneity to areas of particular promises or concerns as the field of PM moves forward.

Methods: A systemic search was performed in Embase, Medline, EconLit, and CRD databases for studies published between 2011 and 2021 on cost-effectiveness analysis (CEA) of PM interventions. Based on a willingness-to-pay threshold of one-time GDP *per capita* of each study country, the net monetary benefit (NMB) of PM was pooled using random-effects meta-analyses. Sources of heterogeneity and study biases were examined using random-effects meta-regressions, jackknife sensitivity analysis, and the biases in economic studies checklist.

Results: Among the 275 unique CEAs of PM, publicly sponsored studies found neither genetic testing nor gene therapy cost-effective in general, which was contradictory to studies funded by commercial entities and early stage evaluations. Evidence of PM being cost-effective was concentrated in a genetic test for screening, diagnosis, or as companion diagnostics (pooled NMBs, \$48,152, \$8,869, \$5,693, $p < 0.001$), in the form of multigene panel testing (pooled NMBs = \$31,026, $p < 0.001$), which only applied to a few disease areas such as cancer and high-income countries. Incremental effectiveness was an essential value driver for varied genetic tests but not gene therapy.

Conclusion: Precision medicine's value for money across application types and contexts was difficult to conclude from published studies, which might be subject to systematic bias. The conducting and reporting of CEA of PM should be locally based and standardized for meaningful comparisons.

KEYWORDS

precision medicine, medical genetics, economic evaluation, value for money, systematic review, meta-analysis, cost effectiveness

Introduction

Precision medicine (PM) is a novel medical approach that tailors intervention decisions based on expression profiling of individual phenotypes and genotypes or directly corrects pathogenic gene mutations (1, 2). The rapid evolvement of PM technology (3, 4) has led to global efforts of introducing PM into the existing healthcare settings to transform healthcare (5–8). However, the clinical adoption rate of PM remains low (9–13), and due to the lack of knowledge about PM's value for money, the incentives among key stakeholders are poorly aligned to catalyze its development and adoption (9, 12, 14).

Cost-effectiveness analysis (CEA) provides a systemic framework to inform such decisions which, over a relevant time horizon of expected PM benefits and within the context of societal willingness-to-pay thresholds (WTP) for such benefits, assesses the cost of an intervention relative to the expected health gains in standard terms, such as quality-adjusted life year (QALY) (15). CEAs are commonly used to inform public and private sectors' reimbursement decisions, clinical guidelines, benefit designs, and price negotiations (16) ("conventional CEA"), and help in decisions regarding product profile development and research priorities at an early clinical cycle ("early CEA") (17). To guide research, practice, and policy related to PM, it is valuable to have a detailed understanding of the CEA literature, focusing on how a reported value is related to contexts and conditions of PM interventions, as well as specifications and potential biases of CEAs. Previous reports have described the general relationship between various characteristics of PMs studied and estimated cost-effectiveness (18, 19). However, previous reports have not formally assessed this literature using meta-analytic approaches.

This study aimed to quantify heterogeneity in the value for money of PM by pooling the net monetary benefits (NMBs) across the types of PM application [(1) screening for genetic conditions that predispose to disease, (2) early diagnosis, (3) prediction of disease progression, (4) companion diagnostic for targeting drug selection, and (5) gene therapy for established condition], as well as other contexts related to PM technology, disease domain, clinical stage, country capacity, and funder types. A secondary objective was to quantify sources of heterogeneity in PM's value for money in the areas of particular promise or concern as the field of PM moves forward.

Methods

The review was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (20), and the protocol was recently published (PROSPERO: 2021 CRD42021272956) (21).

Search strategy and study selection

We conducted the systematic search and study selection using the Covidence platform®. Embase, MEDLINE Ovid, EconLit, CRD, and Web of Science databases were searched to identify relevant studies published between January 1, 2011 and July 8, 2021, limited to studies published in or translated into English. In addition, we searched gray literature from reimbursement dossiers of several HTA agencies. Appendix 1 presents the details of the search strategies and search results for each database. To satisfy the inclusion criteria, the study had to be original research of cost-effectiveness pertaining to human subjects, reporting costs and either LYs, QALYs, disability-adjusted life years (DALYs), or incremental cost-effectiveness ratios (ICERs), and the intervention of interest had to conform to the working definition of PM (2). Selected studies with overlapped contents were excluded by five independent reviewers.

Data extraction

Data were independently extracted by five reviewers which included characteristics of study (author's name, publication year, geographic region, country-income level, type of funders, and conflict of interest), study population (target population, cascade testing, age, sex, disease areas, and associated prevalence and mortality rates), PM intervention (intervention type, profiling method, developmental stage, clinical pathways, test accuracy, uptake, and treatment compliance), comparators, outcomes (economic and effectiveness parameters, surrogate outcome, data source, and willingness-to-pay thresholds), and modeling (study perspective, time horizon, model type, discount rates, measures of dispersion, and uncertainty). Meanwhile, the risk of bias in the CEAs was assessed using the modified economic evaluations bias (ECOBIAS) checklist (22), which assesses sources of heterogeneity and bias in the overall structure and model of economic evaluations.

Data harmonization and statistical analysis

The statistical analysis was performed using Stata MP version 17. The primary outcome was the net monetary benefit (NMB), which measures the difference between a monetized equivalent of incremental effectiveness (i.e., multiplied by a WTP threshold) and the incremental cost of new technology. Based on the central limit theorem, NBM is distributed normally and thus commonly used for quantitative analysis of CEAs (19, 23, 24). Although the standard practice typically uses nationally specific WTP thresholds, to enable global comparison that involves low- and middle-income countries (LMICs), we followed the recommendation of the World Health Organization (WHO) and World Bank (25) that defined the WTP threshold as the one-time national gross domestic product (GDP) *per capita* as of the study year. To standardize costing data, all NMBs were first inflated to the 2020 currency of that study country and then converted to 2020 USD ($) according to the consumer price index and exchange rate from the World Bank (26).

Following the latest guideline for data harmonization in meta-analyses of CEAs (27), we prepared NMB data, with details and the published protocol described in Appendix 2 (21). Through data harmonization, the NMB and its variance were consistently calculated by comparing PM to a conventional intervention strategy. Based on the COMER methodology (28), we performed a random-effects

meta-analysis to calculate weighted-pooled summary estimates of NMB using the DerSimonian and Laird (DL) method (29).

$$Pooled\ NMB_r = \frac{\sum_{i=1}^{N} w_i NMB_i}{\sum_{i=1}^{N} w_i + \gamma^2}, \text{ where } \aleph^2 = \frac{Q-(N-1)}{\sum_{i=1}^{N} w_i - \frac{\sum_{i=1}^{N} w_i^2}{\sum_{i=1}^{N} w_i}}$$

where w_i refers to the inverse of variance. Heterogeneity was tested using the Cochran Q test and $I(2)$ statistics (30), with $I(2) = 25$–74% indicating moderate heterogeneity and $I(2) \geq 75\%$ indicating high heterogeneity.

Subgroup analyses were performed in ≥two datapoints to investigate the context-specific value for money of PM. We estimated the weighted-pooled NMB by subgroups, namely, PM applications, technology [single-gene profiling, multigene panel, whole-genome sequencing (GS), and whole-exome sequencing (ES)], clinical stage (first-clinical-use vs. market access), 16 major disease areas defined by International Disease Classification diagnosis codes, version 10 (ICD-10) (31), WHO region (32), World Bank country-income level (*per capita* Gross National Income in 2020 USD when most information was available) (33), and funder type (public vs. non-profit private, for-profit private, and mixed or unspecified funding sources).

To assess the robustness and conclusiveness of pooled NMB findings, the jackknife sensitivity analysis was performed for each abovementioned subgroup, which omitted one study at a time and repeated the meta-analysis in the rest of the studies (34). This examined whether pooled NMB was consistent across the studies or excessively affected by any influential CEAs.

Following expert recommendation, publication bias was assessed using funnel plots and Egger's test (27). A funnel plot put NMB estimates on the *x*-axis against the quantified uncertainty interval on the *y*-axis. Egger's test assessed whether the funnel was symmetrical, or there was heterogeneity and/or missing studies.

To identify and quantify sources of heterogeneity in the pooled NMB of each PM type, first, we ranked the frequency of the most sensitive parameters to ICER that were reported in the sensitivity analyses of CEAs. Second, we performed univariate random-effects meta-regressions to examine the impact of 19 influencing factors that explain NMB heterogeneity due to study year, target population (age, sex, disease incidence rate, and use of cascade testing), and intervention characteristics (PM cost, incremental effectiveness, integrations of test uptake, test accuracy, and treatment compliance) and that explain value bias as a function of methods (study perspective, time horizon, model type, respective sources of cost and effectiveness data, any use of surrogate outcome, % of "yes" answers in overall ECOBIAS assessment, % "yes" answers in model-specific ECOBIAS assessment, and any conflict of interest). Third, because many covariates were found to be associated with NMB in the univariate meta-regressions ($p < 0{\cdot}05$), we used a generalized Lasso approach with 10-fold cross validation to select essential covariates to be included in the multivariate meta-regression (the best-fitting model) (35). Finally, essential covariates were included in a multivariate, random-effect meta-regression (35) to quantify the impact of essential value drivers on the NMB of each PM type. Of note, we compared three random-effect meta-regression models, namely REML, DL, and empirical Bayes, and selected the model that yielded the greatest reduction in between-study heterogeneity [$\tau(2)$] of NMBs.

Results

Literature search and study characteristic

The literature search initially identified 5,187 articles. The final analysis included 275 unique CEAs with 463 cost-effectiveness estimates of varied PM applications because one CEA may include multiple test-treatment strategies, comparators, and settings (Figure 1, Flowchart of literature search and selection; Appendix 3, Full list of included studies).

Table 1 presents the study characteristics. Appendix 4 provides more details. Among the 238 CEAs on genetic testing and 37 CEAs on gene therapy, most were performed in high-income countries, in Western countries and applied to cancer. The median unit cost ranged between $220 and $3,091 for genetic tests and was $321,268 for gene therapy. The median ΔQALY was the lowest in the prognostic test compared to other test types (0.07 vs. 0.23–0.73) and the highest (3.83) in gene therapy. The pattern of risk of bias was persistent across varied PM application types, mainly focusing on narrow perspective, cost measurement omission, intermittent data collection, double counting, limited sensitivity analysis, and limited scope (Appendix 5).

Context-level variations in PM's value for money

Genetic testing

High heterogeneity was detected from the meta-analysis of 369 cost-effectiveness estimates ($I(2) = 100\%$). By clinical applications, pooled NMBs descended from genetic tests use for screening ($48,152 [95% CI 40,725–55,579]), diagnosis ($8,869 [7,570—10,168]), companion diagnostic for targeted therapy ($5,693 [4,548—6,839]), and to not being significantly greater than 0 for prognostic tests ($2,694 [−601 to 5,988], $p = 0.11$). By profiling technology, multigene panel testing had higher pooled NMB than single gene testing and GS ($31,026 [25,602–36,449] vs. $3,893[3,058–4,727] and $2,429[1,886–2,972], respectively), whereas the pooled NMB of ES was not significantly positive ($p = 1.00$; Figure 2A).

Within each test type, only certain disease areas showed evidence of cost-effectiveness in general. Genetic tests had positive pooled NMBs when used for screening in endocrine and metabolic diseases (especially familial hypercholesterolemia) and cancer, in particular breast cancer ($96,018, $57,889, and $187,000, respectively), for diagnosis in Barrett's esophagus (a pre-malignant digestive condition) and cancer, most commonly thyroid cancer ($58,975, $8,422, and $6,051, respectively), and as a companion diagnostic in chronic infectious diseases (chronic hepatitis C, HIV), gout, and rheumatoid arthritis ($61,333, $4,850, and $4,173, respectively). Nonetheless, the pooled NMBs of the prognostic test were not statistically positive in varied disease areas (Figure 3A).

Gene therapy

In the meta-analysis of 56 cost-effectiveness estimates, the pooled NMB of gene therapy was not significantly greater than 0 in a variety of contexts (Figures 2B, 3B).

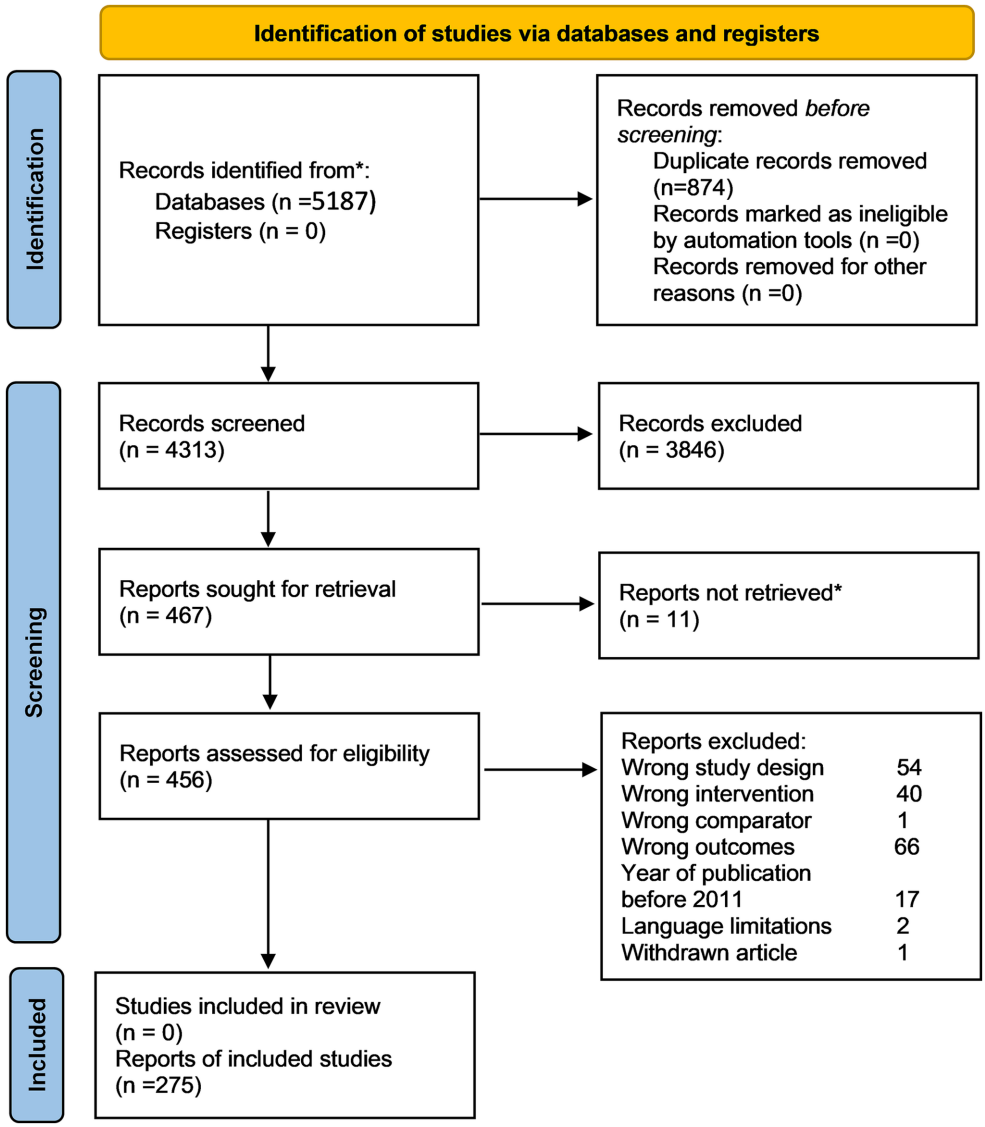


FIGURE 1
PRISMA 2020 flow diagram of systematic search and selection. Numbers refer to unique study records, not datasets, except where otherwise indicated.

System-level variations in PM's value for money

At the structural level, for both genetic tests and gene therapy, commercially funded studies yielded high pooled NMBs, whereas publicly sponsored studies found no evidence of PM being cost-effective in general (Figures 2A,B). Early CEAs also reported a higher pooled NMB compared to conventional CEAs both in genetic tests ($26,009 vs. $16,215; Figure 2A) and gene therapy ($1,830,000 vs. $0 [insignificant value]; Figure 2B).

At the country level, genetic tests had positive pooled NMBs in studies from America ($44,972), Europe ($5,005), and high-income countries ($18,930), whereas an inconclusive value in Western Pacific ($p = 0.72$) and middle-income countries ($p = 0{\cdot}87$ and $0{\cdot}44$). Gene therapy had a negative pooled NMB in studies in Europe (−$588,000) and an inconclusive value in the Americas, Eastern Mediterranean (Qatar), and Western Pacific ($p = 0.57$, 0.60, and 0.88, respectively).

Consistency, robustness, and publication bias of PM's value for money

In the jackknife sensitivity analysis, cost-effectiveness findings were valid and consistent in the above-described subgroups, i.e., both the pooled NMB and the corresponding 95% CI remained in the original position and direction regardless of the omission of any single datapoint (Appendix 6).

As seen by the asymmetry on the funnel plots (Appendix 7), publication bias was present in pooled NMBs of genetic tests in general (Egger's test, coefficient = −0.75, SE = 0.10, $p < 0.001$), particularly in screening, diagnosis, and companion diagnostics

10.3389/fpubh.2023.1151504

TABLE 1 General and economic characteristics of cost-effectiveness analyses reporting precision medicine interventions.

Characteristic	Screening test (*N* = 52)	Diagnostic test (*N* = 27)	Prognostic test (*N* = 53)	Companion test (*N* = 106)	Gene therapy (*N* = 37)
PM Unit Cost, Median (IQR)	385 (147–1,204)	1,059 (424–3,696)	3,091 (754–3,750)	220 (108–439)	321,268 (4,051–607,118)
WHO region, *n* (%)					
African Region (AFR)	0 (0%)	0 (0%)	0 (0%)	1 (0.9%)	0 (0%)
Region of the Americas (AMR)	24 (46%)	15 (56%)	26 (49%)	46 (43%)	21 (57%)
South-East Asian Region (SEAR)	0 (0%)	0 (0%)	0 (0%)	10 (9.4%)	1 (2.7%)
European Region (EUR)	17 (33%)	10 (37%)	22 (42%)	23 (22%)	7 (19%)
Eastern Mediterranean Region (EMR)	1 (1.9%)	0 (0%)	0 (0%)	0 (0%)	1 (2.7%)
Western Pacific Region (WPR)	10 (19%)	2 (7.4%)	5 (9.4%)	26 (25%)	7 (19%)
Study perspective, *n* (%)					
Societal	8 (15%)	5 (19%)	7 (13%)	21 (20%)	5 (14%)
Healthcare	41 (79%)	19 (70%)	42 (79%)	79 (75%)	30 (81%)
Other (e.g., patient perspective)	3 (5.8%)	3 (11%)	4 (7.5%)	6 (5.7%)	2 (5.4%)
Effectiveness outcomes, *n* (%)					
QALYs	47 (90%)	26 (96%)	50 (94%)	104 (98%)	37 (100%)
Life years	5 (9.6%)	1 (3.7%)	3 (5.7%)	2 (1.9%)	0 (0%)
CEA type by PM stage, *n* (%)					
Early CEA to guide R&D	12 (23%)	4 (15%)	12 (23%)	31 (29%)	10 (27%)
Conventional CEA to inform reimbursement	40 (77%)	23 (85%)	41 (77%)	75 (71%)	27 (73%)
Time horizon, *n* (%)					
Short term ($0 < T \leq 3$ years)	1 (1.9%)	3 (11%)	1 (1.9%)	25 (24%)	2 (5.4%)
Intermediate ($3 < T \leq 10$ years)	4 (7.7%)	3 (11%)	16 (30%)	19 (18%)	6 (16%)
Long term ($10 < T \leq 30$ years)	3 (5.8%)	5 (19%)	6 (11%)	11 (10%)	4 (11%)
Lifetime ($T > 30$ years)	43 (83%)	16 (59%)	29 (55%)	49 (46%)	24 (65%)
Not reported	1 (1.9%)	0 (0%)	1 (1.9%)	2 (1.9%)	1 (2.7%)
Conclusion, *n* (%)					
Cost-effective/cost-saving	39 (75%)	21 (78%)	36 (68%)	66 (62%)	23 (62%)
Not cost-effective	10 (19%)	5 (19%)	11 (21%)	31 (29%)	10 (27%)
Inconclusive	3 (5.8%)	1 (3.7%)	6 (11%)	9 (8.5%)	4 (11%)

CEA, Cost-effectiveness analysis; IQR, Inter-quartile range; LY, Life year; QALY, Quality-adjusted life year; NA, Not applicable; QALY, Quality-adjusted life year.

(Egger's test, all value of *p*s < 0.05), whereas there was no evidence of publication bias in pooled NMBs of prognostic tests and gene therapy (Egger's test, $p = 0.296$ and 0.608, respectively).

Sources of heterogeneity in PM's value for money

The ICERs of varied genetic tests but not gene therapy were highly sensitive to disease progression rate and test cost; the ICERs of diagnostic, prognostic, companion tests, and gene therapy but not screening tests were highly sensitive to treatment cost and effectiveness; and the ICERs of screening and diagnostic tests but not prognostic or companion tests were highly sensitive to test accuracy (Appendix 9).

In the univariate meta-regressions of NMBs of genetic tests (Appendix 8), 18 out of 19 selected covariates were significantly associated with NMBs of studies of each test type. Multivariate meta-regression results based on Lasso-selected essential features are presented in Table 2. Overall, 97.2% of variability [i.e., $R(2)$] in

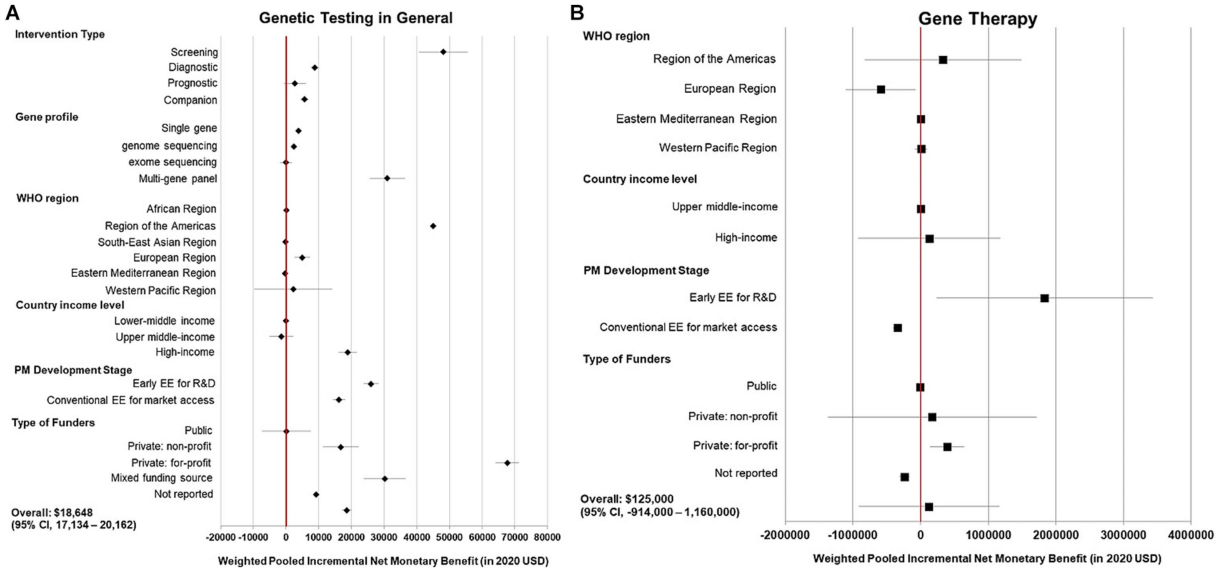


FIGURE 2
Summary forest plot showing the weighted-pooled summary estimates of incremental net monetary benefit of precision medicine. **(A)** Left panel, genetic testing in general; **(B)** Right panel, gene therapy. The error bars show the 95% confidence interval. The red vertical line marks the border for significance.

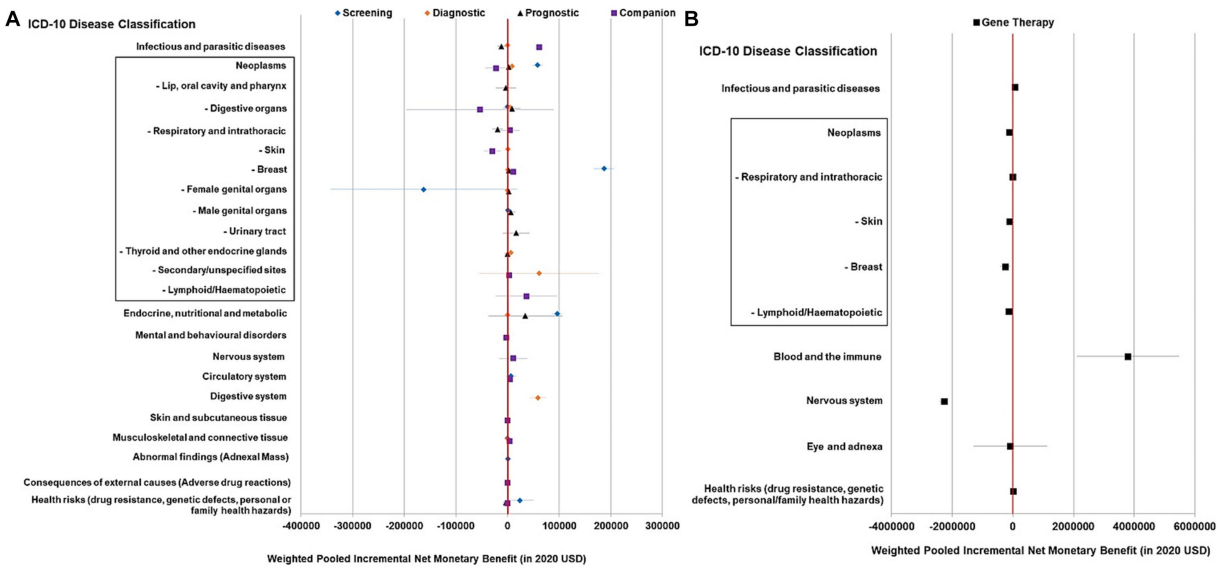


FIGURE 3
Summary forest plot showing the weighted-pooled summary estimates (in ≥ two datapoints) of incremental net monetary benefit of precision medicine across major ICD disease domains. **(A)** Genetic testing for different purposes; **(B)** Gene therapy. The error bars show the 95% confidence interval. The box shows neoplasm/cancer and detailed sub-categories. The red vertical line marks the border for significance.

screening tests' NMBs was explained by incremental effectiveness ($p < 0.001$) and target age ($p = 0.32$), 95.9% of variability in the diagnostic tests' NMBs was explained by incremental effectiveness, target sex, source of cost data, model type, and overall study bias (all had $p < 0.001$), and 48.5% of the variability in the prognostic tests' NMBs was explained by incremental effectiveness ($p < 0.001$), target sex ($p = 0.07$), study perspective ($p < 0.001$), test accuracy ($p = 0.26$), treatment compliance ($p = 0.001$), and publication year ($p = 0.33$), whereas the only essential predictor of the companion tests' NMBs was incremental effectiveness ($p < 0.001$) when treatment cost was absent, explaining 11.8% of the variability. Test cost was not identified as an essential value driver for any genetic test type.

In particular, one extra unit of incremental effectiveness was associated with a marginal increase in NMB of $60,181 (95% CI 59,752–60,609) for the screening test, $41,943 (95% CI 40,381–43,504) for the diagnostic test, $24,515 (95% CI 20,581–28,450) for the prognostic test, and $27,375 (95% CI 26,496–28,255) for the companion diagnostic test.

TABLE 2 Parameter estimates from multivariate meta-regression model on the net monetary benefit of genetic testing.

	Screening		Diagnostic		Prognostic		Companion	
Essential risk factors	coefficient (95% CI)	*p* value	coefficient (95% CI)	*p* value	coefficient (95% CI)	*p* value	coefficient (95% CI)	*p* value
Incremental QALY/LY	60,181 (59,752, 60,609)	**<0·001***	41,943 (40,381, 43,504)	**<0·001**	24,515 (20,581, 28,450)	**<0·001**	27,375 (26,496, 28,255)	**<0·001**
Year of publication					−490 (−1,477, 496)	0·33		
Target age		0·317						
Adult	(Reference)							
Pediatric	−2,898 (−9,329, 3,533)	0·377						
All ages/not specified	−2,931 (−7,286, 1,424)	0·187						
Target sex				**<0·001**		0·068		
Mixed-sex			(Reference)		(Reference)			
All-male			−239,149 (−1,188,713, 710,415)	0·622	10,650 (−8, 21,308)	**0·05**		
All-female			4,161 (2,371, 5,951)	**<0·001**	6,004 (−326, 12,334)	0·063		
Perspective adopted						**0·0002**		
Social					(Reference)			
Healthcare					−15,259 (−22,434, −8,084)	**<0·001**		
Other (e.g., patient perspective)					−14,298 (−25,778, −2,819)	**0·015**		
Type of analysis used for model				**<0·001**				
Decision tree model			(Reference)					
Markov model			−16,034 (−18,868, −13,200)	**<0·001**				
Hybrid model (Decision tree + Markov)			−23,880 (−26,425, −21,335)	**<0·001**				
Discrete event simulation			−18,942 (−22,109, −15,775)	**<0·001**				
Test accuracy integrated					4,877 (−3,527, 13,281)	0·255		
Treatment compliance integrated					14,970 (6,150, 23,789)	**0.001**		
Source of cost data				**<0.001**				
Primary data collected			(Reference)					
Other studies (same setting)			−21,507 (−24,773, −18,241)	**<0.001**				
Secondary sources (same setting)			−2,188 (−4,923, 548)	0.117				
Other studies (other settings)			4,380 (1,607, 7,154)	**0.002**				
% Yes among all ECOBIAS variables, per 1% increase			699 (580, 818)	**<0.001**				

*Bold characters represent a significant effect of a predictor at an alpha level of 0.05. CI, Confidence interval; QALY, Quality-adjusted life year; LY, Life year. Shaded areas indicate that the variable was not essential in predicting the NMB of the corresponding genetic test type.

In the univariate meta-regressions of gene therapy, treatment cost, study perspective, and target patient sex were significant value drivers ($p < 0.001$ for all), but incremental effectiveness barely explained any NMB variability [$R(2) = 0\%$, $p = 0.79$; Appendix 8].

Discussion

In this systematic review and meta-analysis of 275 CEAs on PM published during 2011–2021, the value for money of genetic tests was highly context-specific: While genetic tests appeared cost-effective for screening, diagnosis, or companion diagnosis, such evidence was mainly based on established profiling methods and treatments, well-studied disease indications, and from high-income countries. Evidence in new technologies (e.g., ES and gene therapy) and LMICs remained scarce and inconclusive. Incremental effectiveness and target population but not test cost were the essential drivers of value for money of varied genetic tests. Importantly, studies funded by public agencies generally found NMBs of PM to be not significantly greater than 0, whereas commercially funded and/or early stage studies consistently support PM as cost-effective.

Our findings were generally in line with previous literature. Kasztura et al. (18) reviewed 83 economics studies (2014–2017) on PM and concluded that most previous reviews found inconclusive evidence regarding PM's cost-effectiveness. Vellekoop et al. (19) explored heterogeneity in NMBs of 128 CEAs (2009–2019) on PM. The medians of ΔQALY, Δcost, and NMB of our study were comparable to results of Vellekoop et al. (19) (0.05 vs. 0.03, $445 vs. $575, and $135 vs. $18, respectively). The study by Vellekoop et al. (19) found gene therapies barely cost-effective in general whereas industry sponsorship was positively associated with cost-effectiveness, and our results confirmed both. As an update and extension, we quantified sources of heterogeneity in PM's value for money on an extensive collection of covariates which, for the first time, enabled in-depth investigation into heterogeneity by application type across disease areas, technologies, clinical stages, as well as sources of heterogeneity related to intervention characteristics, model specifications, and study biases.

Across clinical applications, genetic tests reported differential value for money. Of note, test cost was similar across genetic test types and had no major influence on their NMBs. However, one unit increase in ΔQALY would lead to 2–3 times higher ΔNMB if use for screening and diagnosis compared to prognosis or companion diagnostics, which indicated that PM-enabled early intervention (through risk detection or early diagnosis) was more efficient than PM-enabled treatment stratification (by predicted clinical risk or treatment response) in controlling the costs of disease management in general. In support of this, only screening and diagnostic tests appeared as cost-effective in cancer in published studies, whereas prognostic and companion tests were as plentiful in number but appeared to not be cost-effective. In particular, the prognostic test typically stratifies severe subgroups to advanced treatment which may be still patented and costly, rendering in not-cost-effective profiles in general. Furthermore, the cost-effectiveness of the same type of genetic test varied across disease areas probably because it was largely dependent on incremental effectiveness, which can explain the substantial value difference of genetic screening in breast cancer (a genetic test was used for primary screening) vs. cervical cancer (a genetic test was used as an add-on to pap smear screening). Nonetheless, what is subject to change is PM's inconclusive cost-effectiveness profiles in new innovations, new disease indications, and new markets. Over time, the costs of new PM innovation (in particular gene therapy which on average costs $321,268 per patient) can reduce substantially when the scale and scope of production increases, and evidence in new indications and new markets can accumulate. These may render currently not-cost-effective PM interventions to become good value for money in the future.

Our study revealed significant systematic biases. The substantial discrepancies in PM's value for money between early and conventional CEAs, and between commercially funded and publicly sponsored CEAs, can be related to study manipulation as a result of overambition or over-optimism from the R&D community and commercial entities, especially at an early stage when best guesses were commonly adopted, or publication bias such that positive results were more likely to be submitted for publication. For instance, study perspective was found to be an essential value driver of the prognostic genetic test and gene therapy but not of genetic tests used for screening, diagnosis, or companion diagnostics. This pattern could indicate a greater share of analyses leveraging societal perspectives for interventions that were relatively less cost-effective from a healthcare system's perspective. Therefore, our study supports the call from a recent perspective in Nature Reviews (36) that a reference case should be developed to standardize the evaluation and report on the economic impact of PM. For this reason, we are conducting an in-depth analysis of methodological variations that can lead to the development of a reference case for PM evaluation. The results will be published in a separate study.

This study has several limitations. First, it was impossible to capture all sources of heterogeneity due to systemic differences in health service utilization across settings. Second, we excluded non-English publications. Third, using the same WTP for LYs as for QALYs or DALYs may be inappropriate, but over 96% of included studies measured QALYs. Fourth, we were unable to extract treatment cost from the complex, stratified, and/or changing treatment regimens in many studies. Last but not least, the cost-effectiveness findings did not apply to LMICs because no data were available from low-income countries and the studies from middle-income countries provided no support for the cost-effectiveness of PM in general.

To conclude, a large body of evidence suggests that the value for money of PM applications is concentrated in established technologies, disease domains, and markets, which is mainly influenced by incremental effectiveness in favor of early intervention over treatment stratification at diseased stages. It takes time for PM in new innovations, new indications, and new markets to accumulate evidence to affirm its value of money. Moreover, current CEAs of PM are prone to study manipulation and systematic bias. Thus, it is difficult to make an overall conclusion on PM's value for money across application types and disease areas. To enable meaningful comparisons for truly informed decision-making, policymakers and stakeholders should conduct local studies, with appropriate consensus approaches to standardize the conducting and reporting of CEA of PM.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the data are currently available upon request. The data will be later deposited in a central depository in National University of Singapore, Saw Swee Hock School of Public Health, for public access. Requests to access these datasets should be directed to wenjiach@nus.edu.sg.

Author contributions

WC and YT conceptualized the study and identified the research question and method of this study. NW, YW, YZ, DB, and TP performed literature search and data extraction. WC and NW designed the data analysis plan and completed the analysis. SV and DM provided critical policy feedback and supported the design, method, and search strategy of study. WC, YW, SV, DM, and YT supervised the entire research process with support reviewing the search strategy and the research questions. WC and YZ wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by PRECISE Singapore (MOH-000588-01). Thailand's Health Systems Research Institute also partially funded this study for the development of reference case for economic evaluation on precision medicine for health insurance reimbursement in Thailand (HSRI66-099).

Acknowledgments

The authors thank Thunyarat Anothaisintawee (Mahidol University, Thailand) for her comments on the study design and analysis, Hrishkesh M.A. (Health Intervention and Technology Assessment Program, Ministry of Public Health, Thailand) for his contribution to data extraction, and Tan Tag Han (National University of Singapore High School) and Yam Hong Meng (National University of Singapore High School) for their contributions to data processing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1151504/full#supplementary-material

References

1. Schleidgen S, Klingler C, Bertram T, Rogowski WH, Marckmann G. What is personalized medicine: sharpening a vague term based on a systematic literature review. *BMC Med Ethics*. (2013) 14:55. doi: 10.1186/1472-6939-14-55

2. Vellekoop H, Huygens S, Versteegh M, Szilberhorn L, Zelei T, Nagy B, et al. Guidance for the harmonisation and improvement of economic evaluations of personalised medicine. *PharmacoEconomics*. (2021) 39:771–88. doi: 10.1007/s40273-021-01010-z

3. EMBL-EBI (2022). What is Next Generation DNA Sequencing? | Functional genomics II. Available at: https://www.ebi.ac.uk/training/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/next-generation-sequencing/ (Accessed March 15, 2022).

4. YAN SK, LIU RH, JIN HZ, LIU XR, YE J, SHAN L, et al. "Omics" in pharmaceutical research: overview, applications, challenges, and future perspectives. *Chin J Nat Med*. (2015) 13:3–21. doi: 10.1016/S1875-5364(15)60002-4

5. Precision Health Research, Singapore (PRECISE) (2021). Available at: https://www.npm.sg/ (Accessed August 10, 2021).

6. Thailand Pharmacogenomics Network (2021). Available at: http://www.thailandpg.org/ (Accessed August 10, 2021).

7. Ginsburg GS, Phillips KA. Precision medicine: from science to value. *Health Aff*. (2018) 37:694–701. doi: 10.1377/hlthaff.2017.1624

8. NCBI (2021). Genetic Testing Registry (GTR). Available at: https://www.ncbi.nlm.nih.gov/gtr/ (Accessed August 10, 2021).

9. Davis JC, Furstenthal L, Desai AA, Norris T, Sutaria S, Fleming E, et al. The microeconomics of personalized medicine: today's challenge and tomorrow's promise. *Nat Rev Drug Discov*. (2009) 8:279–86. doi: 10.1038/nrd2825

10. Leapman MS, Wang R, Ma S, Gross CP, Ma X. Regional adoption of commercial gene expression testing for prostate Cancer. *JAMA Oncol*. (2021) 7:52–8. doi: 10.1001/jamaoncol.2020.6086

11. Elverum K, Whitman M. Delivering cellular and gene therapies to patients: solutions for realizing the potential of the next generation of medicine. *Gene Ther*. (2020) 27:537–44. doi: 10.1038/s41434-019-0074-7

12. Virelli CR, Mohiuddin AG, Kennedy JL. Barriers to clinical adoption of pharmacogenomic testing in psychiatry: a critical analysis. *Transl Psychiatry*. (2021) 11:509. doi: 10.1038/s41398-021-01600-7

13. Mrazek DA. Psychiatric pharmacogenomic testing in clinical practice. *Dialogues Clin Neurosci*. (2010) 12:69–76. doi: 10.31887/DCNS.2010.12.1/dmrazek

14. Messner DA, Koay P, Al Naber J, Cook-Deegan R, Majumder M, Javitt G, et al. Barriers to clinical adoption of next-generation sequencing: a policy Delphi panel's solutions. *Perinat Med*. (2017) 14:339–54. doi: 10.2217/pme-2016-0104

15. Garber AM, Sculpher MJ. Chapter eight—cost effectiveness and payment policy In: MV Pauly, TG Mcguire and PP Barros, editors. *Handbook of Health Economics*, vol. *2*: North Holland: Elsevier. 471–97.

16. Pietzsch JB, Paté-Cornell ME. Early technology assessment of new medical devices. *Int J Technol Assess Health Care*. (2008) 24:36–44. doi: 10.1017/S0266462307080051

17. IJzerman MJ, Koffijberg H, Fenwick E, Krahn M. Emerging use of early health technology assessment in medical product development: a scoping review of the literature. *PharmacoEconomics*. (2017) 35:727–40. doi: 10.1007/s40273-017-0509-1

18. Kasztura M, Richard A, Bempong NE, Loncar D, Flahault A. Cost-effectiveness of precision medicine: a scoping review. *Int J Public Health*. (2019) 64:1261–71. doi: 10.1007/s00038-019-01298-x

19. Vellekoop H, Versteegh M, Huygens S, Corro Ramos I, Szilberhorn L, Zelei T, et al. The net benefit of personalized medicine: a systematic literature review and regression analysis. *Value Health*. (2022) 25:1428–38. doi: 10.1016/j.jval.2022.01.006

20. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. (2015) 349:g7647. doi: 10.1136/bmj.g7647

21. Chen W, Anothaisintawee T, Butani D, Wang Y, Zemlyanska Y, Wong CBN, et al. Assessing the cost-effectiveness of precision medicine: protocol for a systematic review and meta-analysis. *BMJ Open*. (2022) 12:e057537. doi: 10.1136/bmjopen-2021-057537

22. Adarkwah CC, van Gils PF, Hiligsmann M, Evers SMAA. Risk of bias in model-based economic evaluations: the ECOBIAS checklist. *Expert Rev Pharmacoecon Outcomes Res*. (2016) 16:513–23. doi: 10.1586/14737167.2015.1103185

23. Haider S, Chaikledkaew U, Thavorncharoensap M, Youngkong S, Islam MA, Thakkinstian A. Systematic review and Meta-analysis of cost-effectiveness of rotavirus vaccine in low-income and lower-midle-income countries. *Open Forum Infect Dis*. (2019) 6:ofz117. doi: 10.1093/ofid/ofz117

24. Bagepally BS, Gurav YK, Anothaisintawee T, Youngkong S, Chaikledkaew U, Thakkinstian A. Cost utility of sodium-glucose cotransporter 2 inhibitors in the

treatment of metformin monotherapy failed type 2 diabetes patients: a systematic review and Meta-analysis. *Value Health*. (2019) 22:1458–69. doi: 10.1016/j.jval.2019.09.2750

25. Robinson LA, Hammitt JK, Chang AY, Resch S. Understanding and improving the one and three times GDP per capita cost-effectiveness thresholds. *Health Policy Plan*. (2017) 32:141–5. doi: 10.1093/heapol/czw096

26. World Bank (2021). Inflation, consumer prices (annual %)|Data. Available at: https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG (Accessed August 3, 2021)

27. Bagepally BS, Chaikledkaew U, Chaiyakunapruk N, Attia J, Thakkinstian A. Meta-analysis of economic evaluation studies: data harmonisation and methodological issues. *BMC Health Serv Res*. (2022) 22:202. doi: 10.1186/s12913-022-07595-1

28. Crespo C, Monleon A, Díaz W, Ríos M. Comparative efficiency research (COMER): meta-analysis of cost-effectiveness studies. *BMC Med Res Methodol*. (2014) 14:139. doi: 10.1186/1471-2288-14-139

29. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. (1986) 7:177–88. doi: 10.1016/0197-2456(86)90046-2

30. West SL, Gartlehner G, Mansfield AJ, Poole C, Tant E, Lenfestey N, et al. (2021). Table 7, Summary of common statistical approaches to test for heterogeneity. Published September 2010. Available at: http://www.ncbi.nlm.nih.gov/books/NBK53317/table/ch3.t2/ (Accessed August 3, 2021).

31. World Health Organization (2021). International Classification of Diseases (ICD). Available at: https://www.who.int/standards/classifications/classification-of-diseases (Accessed August 10, 2021).

32. World Health Organization (2021). WHO country-region. Available at: https://www.who.int/countries (Accessed August 10, 2021).

33. World Bank (2021). World Bank Country and Lending Groups. Available at: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (Accessed August 10, 2021).

34. Miller RG. The jackknife-a review. *Biometrika*. (1974) 61:1–15. doi: 10.1093/biomet/61.1.1

35. Ghosh D, Zhu Y, Coffman DL. Penalized regression procedures for variable selection in the potential outcomes framework. *Stat Med*. (2015) 34:1645–58. doi: 10.1002/sim.6433

36. Payne K, Gavan SP, Wright SJ, Thompson AJ. Cost-effectiveness analyses of genetic and genomic diagnostic tests. *Nat Rev Genet*. (2018) 19:235–46. doi: 10.1038/nrg.2017.108

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 07 December 2023
DOI 10.3389/fpubh.2023.1081339

OPEN ACCESS

EDITED BY
Yingying Xu,
Beihang University, China

REVIEWED BY
Peiyao Lu,
Jinan University, China
Jasper Ubels,
German Cancer Research Center (DKFZ), Germany

*CORRESPONDENCE
Zhaoxin Wang
✉ supercell002@sina.com
Tao Yang
✉ yangtfxl@sina.com
Ying Chen
✉ drevonne@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 October 2022
ACCEPTED 21 November 2023
PUBLISHED 07 December 2023

CITATION
Lv Y, Wang Z, Yuan L, Cheng F, Wu H, Wang Z, Yang T and Chen Y (2023) A cost-effectiveness analysis of pre-pregnancy genetic screening for deafness: an empirical study in China. *Front. Public Health* 11:1081339. doi: 10.3389/fpubh.2023.1081339

A cost-effectiveness analysis of pre-pregnancy genetic screening for deafness: an empirical study in China

Yipeng Lv[1†], Zhili Wang[2,3,4†], Ling Yuan[1], Fan Cheng[5], Hao Wu[2,3,4], Zhaoxin Wang[6*], Tao Yang[2,3,4*] and Ying Chen[2,3,4*]

[1]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Department of Otolaryngology-Head and Neck Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3]Ear Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [4]Shanghai Key Laboratory of Translational Medicine on Ear and Nose Diseases, Shanghai, China, [5]Department of Endodontics, Stomatological Hospital and Dental School of Tongji University, Shanghai Engineering Research Center of Tooth Restoration and Regeneration, Shanghai, China, [6]The First Affiliated Hospital, Hainan Medical University, Haikou, Hainan, China

Objectives: This research aims to assess the effectiveness and cost-effectiveness of pre-pregnancy deafness screening policies.

Methods: Married couples from Shanghai, Beijing, and Suzhou in China were enrolled. We conducted high-throughput, pre-pregnancy genetic screenings for deafness in women and their partners. We compared the cost-effectiveness of deafness genetic screening with the status quo. The two-step screening (wife then partner) and following treatments and interventions were included in the decision tree model. We conducted a cost-effectiveness analysis based on the decrease in deaf newborns, healthy newborn births, and cost-utility analysis of pre-pregnancy deafness genetic screening separately. Cost, utility, and probability data used in the three models were collected from a survey combined with literature and expert consultants. A 5% discount rate and a series of one-way sensitivity analyses along with a Monte Carlo simulation were used to test the reliability of this research.

Results: Between Jan 1, 2019, and Dec 31, 2021, we recruited 6,200 females and 540 male spouses from community health service centers in Shanghai, Beijing, and Suzhou. The incremental cost-effectiveness ratio (ICER) for reducing deaf newborn births was USD 32,656 per case and USD 1,203,926 per case for increasing one healthy newborn birth. This gap exists because of the overall decrease of newborn births. From the perspective of the whole society, deafness genetic screening is not cost-effective for reducing the overall quality-adjusted life years (QALY) in the population.

Discussion: Pre-pregnancy genetic testing is effective in decreasing the occurrence of congenital deafness. It is a cost-saving measure when compared with the costs of future medical expenditure and income loss for the affected families. However, such screening and preventive avoidance of pregnancy will decrease the population size and QALY. Only post-screening ART with PGT was shown to increase the birth of healthy newborns. Focusing on key groups such as premature births or consanguineous couples may improve the societal effects of screening.

KEYWORDS

pre-pregnancy genetic screening, deafness screening, cost-effectiveness, cost utility analysis, public health intervention, health in China

Introduction

Congenital deafness has many effects on the quality life of the affected individual and their family and impacts their society. Without timely diagnosis and treatment, deafness can impair language acquisition, mental health, education, work, and income opportunities. The incidence of congenital deafness is 1–3% worldwide; over 30,000 newborn cases are identified in China every year (1, 2). Approximately 50% of congenital deafness is hereditary (3). Congenital deafness can impair quality of life and reduce quality-adjusted life-years (QALY), increasing the societal burdens of disease. Cochlear implants (CI) and hearing aids (HA) in conjunction with speech therapy can help improve the quality of life and communication skills in hearing-impaired children (4). However, these interventions are not curative and may not completely return the affected individual's quality of life to normal levels.

Genetic factors are responsible for over 50% of hearing loss encountered in neonates and nearly 40% in children (5, 6). Approximately 80% of genetic hearing loss is autosomal recessive; many cases are born from spouses without a family history of congenital or childhood hearing loss. Among these cases, mutations in the deafness genes *GJB2* and *SLC26A4* are the most prevalent in many countries, including China (PMID: 31564438, 30890784).

The main intervention approaches for preventing congenital deafness involve three strategies. Primary prevention involves deafness genetic screening, genetic counseling, and fertility guidance before pregnancy. Secondary prevention is prenatal deafness genetic screening and diagnosis. Tertiary prevention is newborn hearing screening, diagnosis, and intervention with language rehabilitation. Pre-pregnancy deafness genetic screening can identify risks for deafness and allow parents to make informed pregnancy-related choices around their risk of birth to children with hearing loss. Genetic diagnosis of hearing loss can help to avoid unnecessary and costly clinical testing, offer prognostic information, and guide future medical management. The importance of an etiological diagnosis is underlined by the 2014 American College of Medical Genetics and Genomics (ACMG) guidelines for the diagnosis of hearing loss, which recommended that genetic testing should be included in the workup of patients with non-syndromic hearing loss (NSHL) (7).

Currently, the universal newborn hearing screening program (UNHS) has been widely used as a hearing screening program in many countries around the world with otoacoustic emission (OAEs) and automated auditory brain stem response (AABR) technologies (2). Conversely, pre-pregnancy genetic screening strategies, as part of a hearing loss prevention policy, remain underutilized in most countries (7). Next-generation sequencing (NGS) technology has been widely implemented in the genetic diagnosis of hearing loss. However, relatively few countries utilize this technology as part of a national policy for pre-pregnancy deafness screening. Given the limits of health funding, understanding the cost-effectiveness of such a policy in China is critical. Therefore, in this study, we collected cost and effectiveness data to assess the effect, utility, and cost-effectiveness of pre-pregnancy deafness screening policy from the perspective of society and affected families. As the prevention of disability and promotion of a healthy population are also important goals for policymakers, the cost-effectiveness for the overall population was also examined.

Methods

Study design

In this study, we performed a high-throughput, pre-pregnancy genetic screening for deafness in women and their spouses from the general population. We used targeted NGS that covers 45 common mutations in the *GJB2* and *SLC26A4* genes (Supplementary Table S1). We collected information about the epidemiological characteristics associated with these deafness-related genes and conducted a cost-effectiveness analysis with expected reproductive outcomes and corresponding costs. We used a two-step screening strategy. In the first step, women planning pregnancy received genetic screening; if negative, their involvement in the study was marked as complete. If pathogenic recessive mutations in the deafness-related genes were identified, their partners were screened in the second step. Based on the results of the genetic screenings of the couples, families were divided into four different risk categories including high-risk, medium-high risk, medium-low risk, and low-risk which reflected their odds of delivering a newborn with genetic hearing loss. These risk categories and their following treatments after genetic screening were shown as follows:

(1) High-risk families: both husband and wife have biallelic recessive mutations in the same gene, and the newborn is very likely to have genetic deafness. For these families, only pre-pregnancy medical counseling services were provided regarding the likely hearing loss. The couples were free to decide whether or not to give birth based on this information.
(2) Medium-high risk families: one of the spouses has biallelic recessive mutations and the other has a single heterozygous recessive mutation in the same gene. These couples can expect a 50% chance of delivering a child with genetic deafness.
(3) Medium-low risk families: Both spouses have a single heterozygous recessive mutation in the same gene. There is an approximately 25% chance of delivering a child with genetic deafness.
(4) Low-risk families: the woman's pregnancy genetic screening result is negative. The chance of delivering a child with genetic deafness is relatively low.

For medium-risk families (categories 2 and 3) the main follow-up interventions included: (1) choosing not to have children; (2) normal

pregnancy; (3) normal pregnancy with a prenatal amniocentesis (if amniocentesis screening was positive, couples could decide whether or not to terminate the pregnancy according to the local legal and ethical regulations); (4) utilize assisted reproductive technologies (ART) with a preimplantation genetic test (PGT) and proceed to implantation of fertilized eggs with the desired genotypes.

Subjects and public involvement

From 2019 to 2021, 6,200 females and 540 male spouses from the general population were recruited from community health service centers in Shanghai, Beijing, and Suzhou (Figure 1); this study involved research teams from the Shanghai Ninth People's Hospital, the Haidian District Maternal and Child Health Care Hospital, and the Suzhou Science and Technology Town Hospital, respectively. The inclusion criterion for female participants was being aged between 20 and 55 years. The sole exclusion criterion was an inability to provide complete demographic and health information as required by this study. All participants provided informed consent before participation and the study was conducted per approval by the Ethics Committee of Ninth People's Hospital, Shanghai Jiaotong University School of Medicine.

Genomic DNA was extracted from blood samples provided by the participants. Sequences covering the 45 common mutations for deafness (Supplementary Table S1) were captured by a customized capture assay (Fujun Genetics, Shanghai, China) and sequenced on an Illumina NovaSeq 6000-PE150 platform (Illumina, San Diego, CA, USA). Pathogenic mutations were confirmed by Sanger sequencing.

Model generation

The genetic deafness screening was conducted in two steps as previously mentioned. The corresponding treatment and intervention were carried out based on the test result and the couple's decision. A decision tree model was used in the research and we compared the deafness genetic screening strategy with the status quo, which in essence means the absence of any intervention that alters the course of the pregnancy. All possible options available to couples following screening were included in the decision tree (mentioned in the STUDY DESIGN part). The three possible outcomes were the birth of a healthy newborn or a deaf newborn and none (Figure 2). All data such as costs, utilities, the probability of different arms in the model were collected from the survey combined with literature and an expert consultant (Tables 1, 2). We analyzed the cost, effect, and ICER between the screening strategy and status quo. The model was calculated by TreeAge Pro (Healthcare Version, 2022 version; TreeAge Software, Inc., Williamstown, MA).

Cost analysis

We included direct medical costs, direct non-medical costs, and indirect costs. The direct medical costs mainly included the screening cost and the cost of medical care after screening, which included the genetic screening cost [calculated by the fixed cost of equipment and the variable cost of a single test, like alcohol swabs, kits (cassette and buffer), gloves, biosecurity devices, electricity, and water consumption], screening-related medical staff salaries, ART with PGT costs, amniocentesis costs, and maternity expenses. The total cost of genetic screening for deafness was 350.00RMB per person, based on contracted service for next-generation sequencing provided by Fujun Genetics, Shanghai, China. The direct non-medical costs included the promotion costs of the screening project including advocacy meetings with authorities, training, supervision, and monitoring. Because the deafness gene screening was carried out in community health centers close to the participants' homes, we did not include costs for additional space or transportation. Indirect costs mainly refer to productivity loss due to the time spent as part of diagnosis, treatment, and medical examination. The cost is calculated based on the *per capita* disposable income of urban and rural areas in China in 2021, as shown in Table 1.

Furthermore, we expected significant differences in future income and medical expenditure between deaf and healthy children. The literature indicates that 47% of the costs arising from deafness are related to the loss of quality of life and 32% are related to additional health and medical expenditure (13). In addition to the provision of hearing aids and cochlear implants, the medical expenses incurred by deaf newborns and their families also include follow-up hearing and speech rehabilitation, as well as future medical costs due to injuries and other services arising from deafness. As a result, the cost of deafness-related medical care is difficult to measure in detail and can vary widely based on individual situations; unfortunately, comprehensive research in this area is lacking. In this study, for both

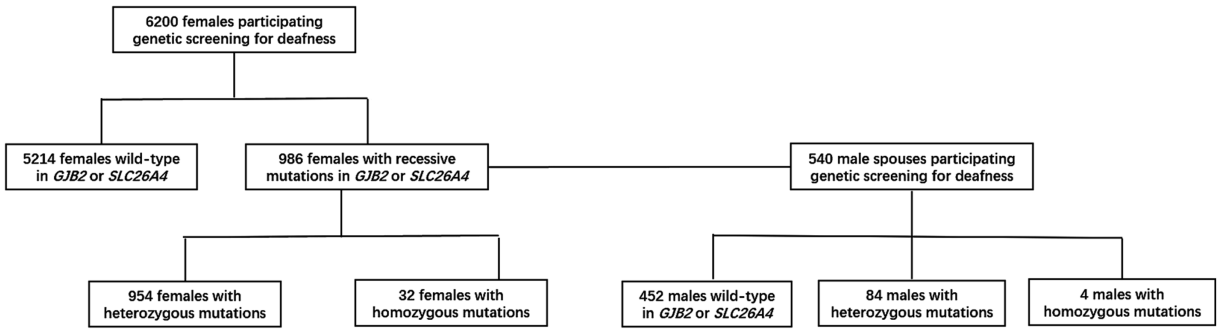


FIGURE 1
The flowchart of objects participating the research and their genetic screening results.

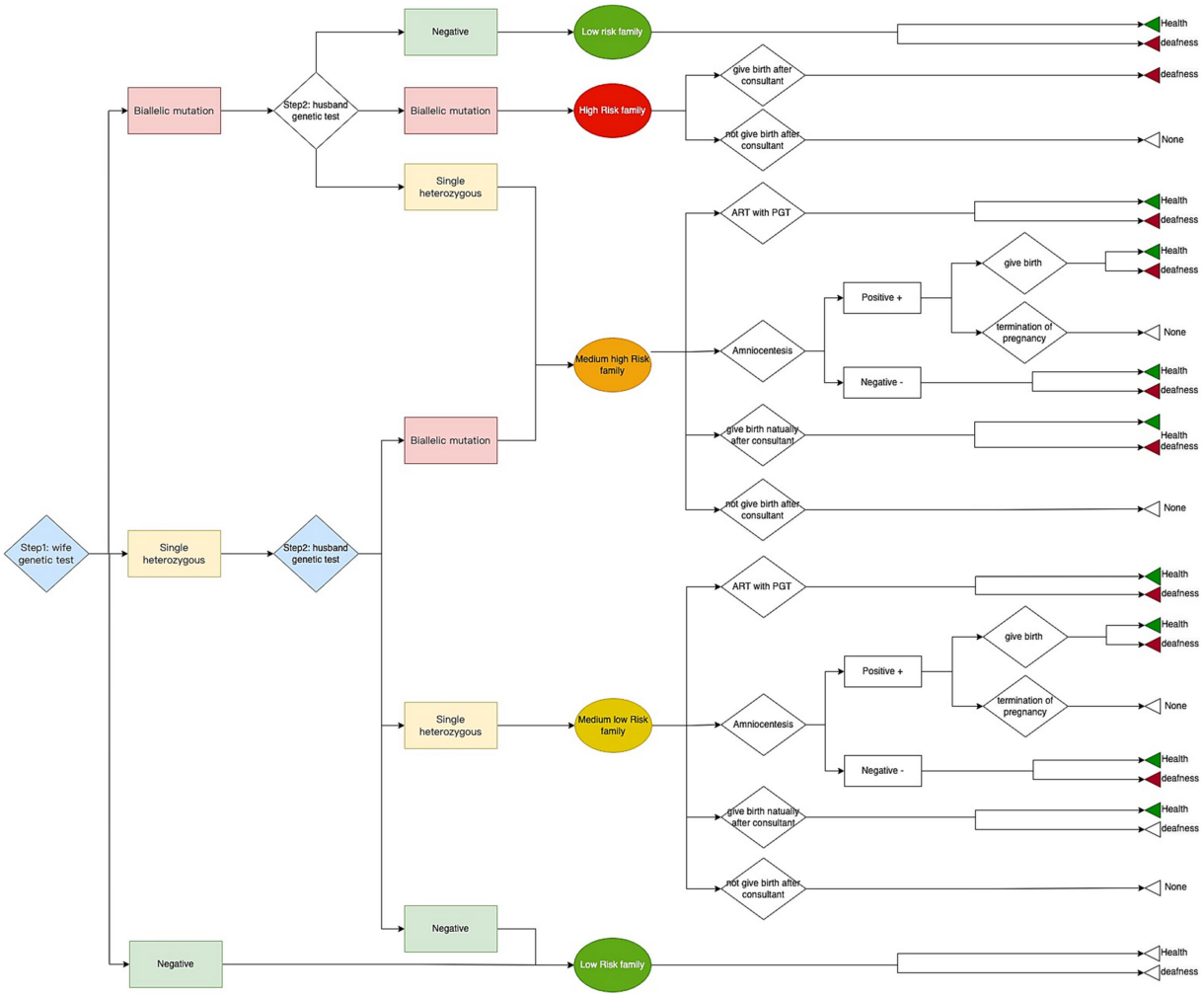


FIGURE 2
Schematic diagram of the decision tree for deafness screening.

deaf and healthy people have medical expenditure and productivity income. This component was calculated based on the annual *per capita* consumption expenditure of disabled households and national resident households from the "*China 2019 National Survey Report on Income of Disabled Households*," which are $402.50/year and $251.50/year, respectively. The annual growth rate of health care expenses is calculated by using the 2021 and 2019 annual *per capita* consumption expenditures of households across China. As a result, the estimated 2021 medical expense for disabled individuals is $505.20/year. The total expenditure in life years is calculated utilizing the Chinese life table and a 5% discount rate as suggested by Guidelines for the Evaluation of Chinese Pharmacoeconomics (Table 1).

Loss of future income

The future labor loss of deaf individuals includes two components: the loss of earning capacity caused by deafness and disability to the individual and to the family members who care for the deaf individual. This cost was estimated by utilizing the *per capita* disposable income of households with disabilities. According to the "*2019 National Survey Report on Income of Households with Disabilities*," the *per capita* disposable incomes of Chinese families with disabilities and national residents in 2018 were $2,040.80/year and $4,213.10/year, respectively. Based on the *per capita* disposable income of Chinese residents in 2021 ($5,243.00) from the National Bureau of Statistics of China, the three-year average annual growth rate was calculated to be 7.56% which resulted in the *per capita* disposable income of Chinese families with disabilities in 2021 to be estimated as $2,992.70. The expenses were calculated according to the Chinese life table with a 5% discount rate, as shown in Table 1 (calculation shown in Supplementary appendix Tables S2, S3).

Monetary costs were adjusted to the average 2021 US Dollar exchange value and are listed in Table 1. A 5% discount rate was used based on recommendations of the China Pharmacoeconomics Committee. In this study, all data such as productivity loss, income, medical expenditure, etc. were translated into the 2021 USD value based on a 5% discount rate. We ran the model in TreeAge Pro (TreeAge software, Williamstown, MA, USA).

TABLE 1 Cost of pre-pregnancy deafness genetic screening and subsequent interventions (2021, USD).

Items	Value	Source
Direct cost ($/case)		
Deafness genetic screening cost	52.20	Collected from survey results
Labor cost for screening	9.00	
Screening project promotion cost	3.00	
Reproductive cost*	651.00	(8, 9)
ART with PGT	11,940.30	Collected from survey results
Amniocentesis	2,686.60	
Indirect cost ($/case)		
Productivity loss of ART with PGT	28.70	(10, 11)
Productivity loss of Amniocentesis	14.40	
Future cost ($/life)†		
Future income of healthy individuals	115,622.00	(11, 12)
Future income of disabled individuals	65,996.00	
Future medical expenditure of health individuals	6,961.40	
Future medical expenditure of disabled individuals	11,142.40	

The average hourly wage in urban and rural China is $3.19 based on the 2021 National Economic and Social Development Statistical Bulletin and the 2018 National Time Use Survey Bulletin issued by the National Bureau of Statistics. Productivity loss of ART and Amniocentesis was 1 and 0.5 days, respectively.
*The reproductive cost was calculated as the cost of natural and cesarean births with their respective proportions in China.
†Productivity loss was calculated according to the utilized days.

Outcome variables and willingness-to-pay

There were two possible pregnancy outcomes: (1) healthy newborns with normal hearing, or (2) newborns with congenital deafness. Due to the complex association between the severity of congenital deafness and the various pathogenic mutations, we did not attempt to subgroup outcomes in this study. The likelihood of outcome was analyzed according to the four family risk subgroup categories.

We built three different models to achieve our research aims: (1) model 1: A cost-effectiveness analysis of taking steps to reduce the birth of deaf newborns. We set the deafness outcome as 1 and the health outcome as 0 and tested the effect of reducing the number of deaf infants. The resulting gap in the model between the screening arm and the status quo arm is the cost-effectiveness of deaf newborns in these two scenarios; (2) model 2: A cost-effectiveness analysis of increasing the birth of healthy newborns. We set the health outcome as 1 and the deafness outcome as 0 and tested the impact of the screening strategy on the birth of healthy newborns; (3) model 3: A cost-utility and policy feasibility analysis of pre-pregnancy deafness genetic screening. The severity of genetic deafness in newborns varies with age. Furthermore, we also built another model to explore the effect of genetic screening on the overall population. This model is not the main focus of this study, so it is shown in Supplementary Appendix. We calculated the health utilization of deaf newborns based on the proportion of different degrees of deafness in China in 2019 from the Chinese global burden of disease data (GBD) multiplied by the utility of different deafness levels from literature, which is 0.91 (14). According to the Chinese life table (15) and the 5% discount rate, the QALY of a healthy person in China is 22.1. Therefore, we set the utility as 22.1 for healthy newborns and 20.1 for deaf newborns (calculation shown in the Supplementary appendix Table S2).

However, the willingness to pay (WTP) range is difficult to translate to the benefit of birth outcomes. There is an evidence gap in fertility-specific WTP guidelines in the field (16). Therefore, for the cost effectiveness analysis (model 1 and 2 to analyze the impact of policy on births of healthy and deaf children), we did not compare it to the regular WTP, but rather compared it to the disease-related opportunity cost. In cost-utility analysis, for the results represented as the QALYs, we compare the ICER for overall QALYs with the whole WTP for the life value. Specifically, China's *per capita* GDP in 2021 was 12,086.00 USD and 22.1 years for QALY for a healthy person's whole life. Consequently, we estimate the WTP in this study as 801,302 which is three times the *per capita* GDP in China for the whole QALY (calculation shown in the Supplementary appendix Tables S2, S3).

Probability of different arms in the model

Table 2 shows the specific values and data sources of the various genotype-intervention pairs and following the various interventions and their probability of corresponding results (the birth of deaf or healthy newborns). The incidence of autosomal recessive pathogenic mutations in the common deafness genes *GJB2* and *SLC26A4* among the Chinese population was collected in the overall cohort of 6,740 individuals. The probability of amniocentesis in the different groups was calculated based on the Mendelian inheritance law. The remaining data were obtained from a literature review combined with expert consultation in the field. For arms with no intervention, the probability of a birth of deaf and healthy newborns was based on Mendelian laws of inheritance.

Sensitivity analysis

Both one-way deterministic and simulated probabilistic sensitivity analyses were conducted to assess the robustness of the main outcomes. For the one-way sensitivity analysis, a change of plus or minus 10% (for prevalence, utility, sensitivity, specificity, compliance, and transition probability from published literature) and plus or minus 20% of the original values for the parameters were used for probability-related data in sensitivity analysis.

For the probabilistic sensitivity analyses, we set the distribution for cost and transition probability. The cost mainly includes the key costs involved in the deafness screening strategy, including the cost of screening, the cost of PGT, and the cost of amniotic fluid examination. As for the transition probability, since our data are derived from the results of real-life research studies, we do not design the distribution of deafness gene proportion, which is mainly the proportion of (1)

10.3389/fpubh.2023.1081339

TABLE 2 The probability of each arm used in the model.

No.	Items	%	Sources
1	Wife biallelic mutation	0.0050	Actual data and expert advice
2	Wife single heterozygous mutation	0.1572	Actual data
3	Wife passes screening (−)	0.8378	
4	Husband Biallelic Mutation	0.0074	
5	Husband single heterozygous mutation	0.1370	
6	Husband passes screening (−)	0.8556	
7	High-risk families choose to give birth naturally	0.1930	(17)
8	High-risk families choose not to have children	0.8070	
9	Medium-risk families choose to take the ART with PGT	0.1774	(18) Combined with expert advice
10	Medium-risk families choose amniocentesis	0.7096	
11	Medium-risk families choose to give birth naturally	0.0565	
12	Medium-risk families choose not to have children	0.0565	
13	Positive amniocentesis in medium-high-risk (+)	0.5000	Calculated by Mendelian laws of inheritance
14	Negative amniocentesis in medium-high-risk family (−)	0.5000	
15	Positive amniocentesis in medium-low-risk family (+)	0.2500	
16	Negative amniocentesis in medium-low-risk family (−)	0.7500	
17	Medium-risk families give birth with positive amniocentesis (+)	0.1930	(17)
18	Medium-risk families end the pregnancy with positive amniocentesis (+)	0.8070	
19	Low-risk families have healthy newborns	0.9990	(6, 18)
20	Low-risk families have deaf newborns	0.0010	
21	Medium-risk families have healthy newborns after ART with PGT	0.9990	(19)
22	Medium-risk families have deaf newborns after ART with PGT	0.0010	
23	Medium-risk families have healthy newborns with positive amniocentesis (+)	0.0010	Expert advice
24	Medium-risk families have deaf newborns with positive amniocentesis (+)	0.9990	
25	Medium-risk families have healthy newborns with negative amniocentesis (−)	0.9990	(20)
26	Medium-risk families have deaf newborns with negative amniocentesis (−)	0.0010	
27	Medium-high-risk families have healthy newborns naturally	0.5000	
28	Medium high-risk families have deaf newborns naturally	0.5000	
29	Medium low-risk families have healthy newborns naturally	0.7500	Expert advice
30	Medium low-risk families have deaf newborns naturally	0.2500	
31	Low-risk families have healthy newborns naturally	0.9990	(6, 21)
32	Low-risk families have deaf newborns naturally	0.0010	

high-risk families (couples who are heterozygous and whose children are born to be with congenital deafness) who choose to continue to give birth naturally; (2) medium-risk families (both parents are heterozygous and whose children have a 50% probability of congenital deafness in natural birth) who choose to take the ART with PGT; (3) medium-risk families choose to give birth naturally with amniocentesis; (4) medium-risk families still choose to give birth naturally without any intervention after being informed of the risks; and (5) medium-risk families choose not to have children. These probabilities are susceptible to subjective influences, and can affect cost-effectiveness outcomes to a large extent. By designing the distribution and conducting sensitivity analyses based on these changeable elements, which have strong influence on the outcomes of a deafness screening strategy, it is possible to better understand the conditions under which such a policy would be more effective, and the patterns of change that are influenced by both environmental and individual subjective variables. A β distribution was applied to prevalence, utilities, and transition probabilities, and a γ distribution was applied to cost parameters.

The random error associated with an estimate for the values was included within a plausible range.

Results

Study population

Between Jan 1, 2019, and Dec 31, 2021, we recruited 6,200 females and 540 male spouses from community health service centers in Shanghai, Beijing, and Suzhou. For women who tested positive, we were able to obtain all (100%) of the associated male partner samples. The distribution of the detected target genes is shown in Table 3.

Measurement of health effects

First, as shown in Table 4, compared with the status quo, each instance of deafness genetic screening was found to reduce the birth of 0.0051 deaf newborns. In other words, 196 families were screened to reduce the birth of deaf newborns by one. Secondly, we looked at the effects on healthy newborns and found that the screening had a small impact on newborns. 7,226 families needed to be screened to increase the birth of healthy newborns by one. Third, our cost-utility analysis showed that genetic screening reduced 0.099 QALY, mainly due to the fewer births of deaf children which resulted in a lower overall QALY.

Cost-effectiveness ratios

In general, two types of costs were included in this study: (1) genetic screening and subsequent medical intervention costs; and (2) future lifetime income and medical expenditure. In part 1 and 2, for the cost-effectiveness analysis of reducing deaf newborns and increasing healthy ones, we included only the costs of screening and the resultant interventions. In part 3, the utility analysis of deafness genetic screening, both the medical expenditures and future income were considered.

As shown in Table 4, the ICER associated with a reduction of deaf newborn births was $32,656.00/case and $1,203,926.00/case for increasing the birth of healthy newborns. From a societal perspective, we found that genetic screening for deafness is not cost-effective for reducing the overall societal QALY. A single screening cost $432.00 and the ICER of utility was $4336.40 /QALY.

Sensitivity analysis

Figure 3 shows the one-way sensitivity analysis of key variables and the ICER distribution in the three models. The variables and results are shown in Supplementary Tables S6–S9.

Figure 3A	Figure 3B
Figure 3C	

In this study, ART with PGT after a positive deafness genetic screening was the only intervention to simultaneously increase healthy newborns and reduce deaf newborns, compared to the status quo. Therefore, we used a one-way sensitivity analysis to test the effects of this intervention. We found that when the percentage of medium-risk families choosing ART with PGT was >17.4%, there were more healthy than deaf infants compared with the status quo (Figure 4).

Category	Incremental ICER	Acceptability curve
Model 1	Figure 4A	Figure 4D
Model 2	Figure 4B	Figure 4E
Model 3	Figure 4C	Figure 4F

Monte Carlo analysis

The incremental cost-effectiveness plot is presented in Figure 5. Here, each dot represents an incremental cost plotted against the incremental effectiveness associated with 10,000 Monte Carlo simulations within our 4 models. The simulation falls under a $801,302/life willingness-to-pay threshold. In model 2, nearly half the simulation result fall under the threshold. In models 3 and 4, the deafness screening strategy is not cost-effective.

Category	ICE Scatterplot
Model 1	Figure 5A
Model 2	Figure 5B
Model 3	Figure 5C

Discussion

To our knowledge, this study involves the largest sample analysis of the cost-effectiveness of pre-pregnancy genetic screening for deafness in the Chinese population. Our findings help inform decision-making about the implementation of a deafness genetic screening policy. The study cohort was recruited from Shanghai, Beijing, and Suzhou, three of the biggest cities in China, and the sample was representative of the urban population.

We found that deafness screening reduced the birth of deaf newborns with an ICER of $32,656 /case. According to the literature, the ICERs of traditional diagnosis and treatment for childhood hearing loss in developing countries, including the cost of cochlear implants and hearing aid installation, are $15,169 /QALY and $15,430/QALY, respectively (14). The costs of screening for other birth abnormalities also provide us with a useful reference. The ICER of preimplantation genetic testing to prevent the transmission of breast and ovarian cancer (BRCA) is $14,242/QALY for BRCA1 and $12,893/QALY for BRCA2 (21). Another study showed that *in vitro* fertilization preimplantation genetic testing for Huntington's disease is associated with 77 more QALYs and a cost savings of $46,394,268. Direct comparisons in outcomes cannot be made between a decrease in ICER associated with deafness and other measures that quantify disability metrics because of their different results (life vs. QALY) and the long impact of deafness on the individual, family, and society. However, if we take $53,807 (the sum of the additional medical expenses and productivity loss of deaf individuals) as the opportunity cost, the ICER decrease associated with reducing deaf newborns shows a better effect compared with the additional medical expenses and productivity loss. Furthermore, significant costs are expected to arise following the birth of a deaf newborn, including additional injuries, medical expenditure, and negative occupational effects due to the disability. A systematic review of 59 studies showed that estimates of the economic cost of productivity loss vary widely, from $1.8 to $194 billion in the United States. Excess medical costs resulting from hearing impairment including audiometric testing and treatment with bilateral hearing aids range from $3.3 to $12.8 billion nationally *per annum* in the United States alone (22).

Through a one-way sensitivity analysis, we found that ICER will increase as couples opt for ART with PGT following an adverse genetic finding. In contrast, as couples chose amniocentesis, the ICER decreased. These outcomes are mainly due to the large cost of ART with PGT compared with the relatively low costs of amniocentesis. However, ART can ensure the birth of healthy newborns and reduce the pain of surgical abortion if the amniocentesis test confirms a genetic disorder in the developing fetus (23). Generally speaking, the promotion of pre-pregnancy deafness genetic screening can effectively reduce congenital deafness and the associated familial and societal burdens of disease. The relatively low screening costs ($64.2 in our case) are acceptable to most patients but their widespread coverage by health insurance should be considered as the test can not only reduce the economic burden for patients but also improve the societal effects of screening policies.

We found that when over 17.4% of couples choose ART with PGT, there was a greater proportion of healthy newborns compared to the status quo. A one-way sensitivity analysis showed that with the increase of medium-risk families choosing ART, the ICER of the screening strategy decreased rapidly. The ICER curve tended to be flat around 0.3 and when the ratio was 1 (where all medium-risk families choose ART) the lowest ICER was $69,994.4/case. Therefore, policymakers would need to attempt to increase the proportion of medium-risk families that opt for ART with PGT after screening to reduce the overall societal costs associated with deafness. Several measures could facilitate this goal including the introduction of medical insurance and community education to promote genetic screening and relevant interventions (24, 25).

TABLE 3 Study population and their deafness genes distribution.

Characteristic	Wild-type *N* (%)	Mono-allelic *N* (%)	Bi-allelic *N* (%)	Total (*N*)
Sex				
Male	452(83.70)	84(15.56)	4(0.74%)	540
Female	5,214(84.1)	954(15.4)	32(0.5)	6,200
Location				
Shanghai	5,141(83.6)	974(15.9)	33(0.5)	6,148
Beijing	446(89.4)	52(10.4)	1(0.2)	499
Suzhou	79(84.9)	12(12.9)	2(2.2)	93

At present, few studies have focused on the impact of medical screening measures on the wider population. The birth rate is regarded as an important indicator of a country's development and affects the country's population size which is one of the most critical factors in policy-making. Therefore, we analyzed the impact of deafness genetic screening policy on the general population in this study. We found that, because screening was more effective in reducing the births of deaf newborns than in promoting the birth of healthy newborns, the overall number of births was reduced. However, reducing the number of deaf newborns will also reduce the cost of social governance (13). In addition, from the families' perspectives, reducing the births of deaf children can significantly reduce intangible costs that were not considered in this study, such as anxiety, distress, and other psychosocial pressures (26, 27).

From a societal perspective, the deafness genetic screening was not able to increase the QALY compared with the status quo. Relatively speaking, in middle-income countries cochlear implants and hearing aids have been shown to increase QALYs by 5.7 and 4.6, respectively, compared with no treatment (14). Three reasons may explain the differences between these and our findings. First, there is a decrease in newborn births due to genetic screening which leads to a decrease in the overall population QALY. Second, screening benefits a wide range of families with unknown risk of deafness severity of their future child. As a result, utility in this study was derived from a Chinese hearing loss burden of disease database which includes hearing impaired newborns with relatively lower disease severity and better utility in general compared with other age groups. This resulted in the limited QALY improvements for the whole population in general. However, individuals who receive HI and CI always have moderate, severe, or profound deafness. Consequently, HI and CI lead to better effects for their target group compared with the pre-pregnancy screening strategy as a widespread form of population-wide screening (28, 29). Thirdly, in this study we only calculated the total utility of deaf newborns derived from a Chinese hearing loss burden of disease database while the lifetime adverse effects of deafness leading to worse utility were ignored; this may have resulted in excessively conservative results (30–32). If the relationship between the deafness genotype and the severity of the hearing loss is better established, genetic screening becomes more effective because it can be oriented towards identifying and avoiding pathogenic mutations that lead to severe or profound deafness. Focusing on high-risk groups such as consanguineous couples can also improve screening outcomes. A WHO report stated that "Genetic hearing loss is encountered more frequently in children born to consanguineous parents (12–15) and consanguineous marriages are a common tradition in many communities across the world."

TABLE 4 ICER of the three different models.

Category	Strategy	Cost	Incr cost	Effectiveness	Incr eff	ICER
Deaf = 1	Status quo	651.00		0.0073		
	Screening	817.60	166.60	0.0022	−0.0051	Dominated[a]
Health = 1	Status quo	651.00		0.9927		
	Screening	817.60	166.60	0.9928	0.0001	1,203,926.40
Utility	Status quo	−107,615.66		22.09		
	Screening	−107,184.25	431.41	21.99	−0.0995	Dominated[b]

The results were negative, indicating that the screening strategy did not improve the QALY of the whole society.

[a]Dominated (i.e., costs more and less effectiveness) VS. Status Quo. The number of deaf newborn births becomes less under the screening strategy.

[b]Dominated (i.e., costs more and less effectiveness) VS. Status Quo. In the cost utility analysis, we compared the impact of the screening strategy on QALY with the current situation.

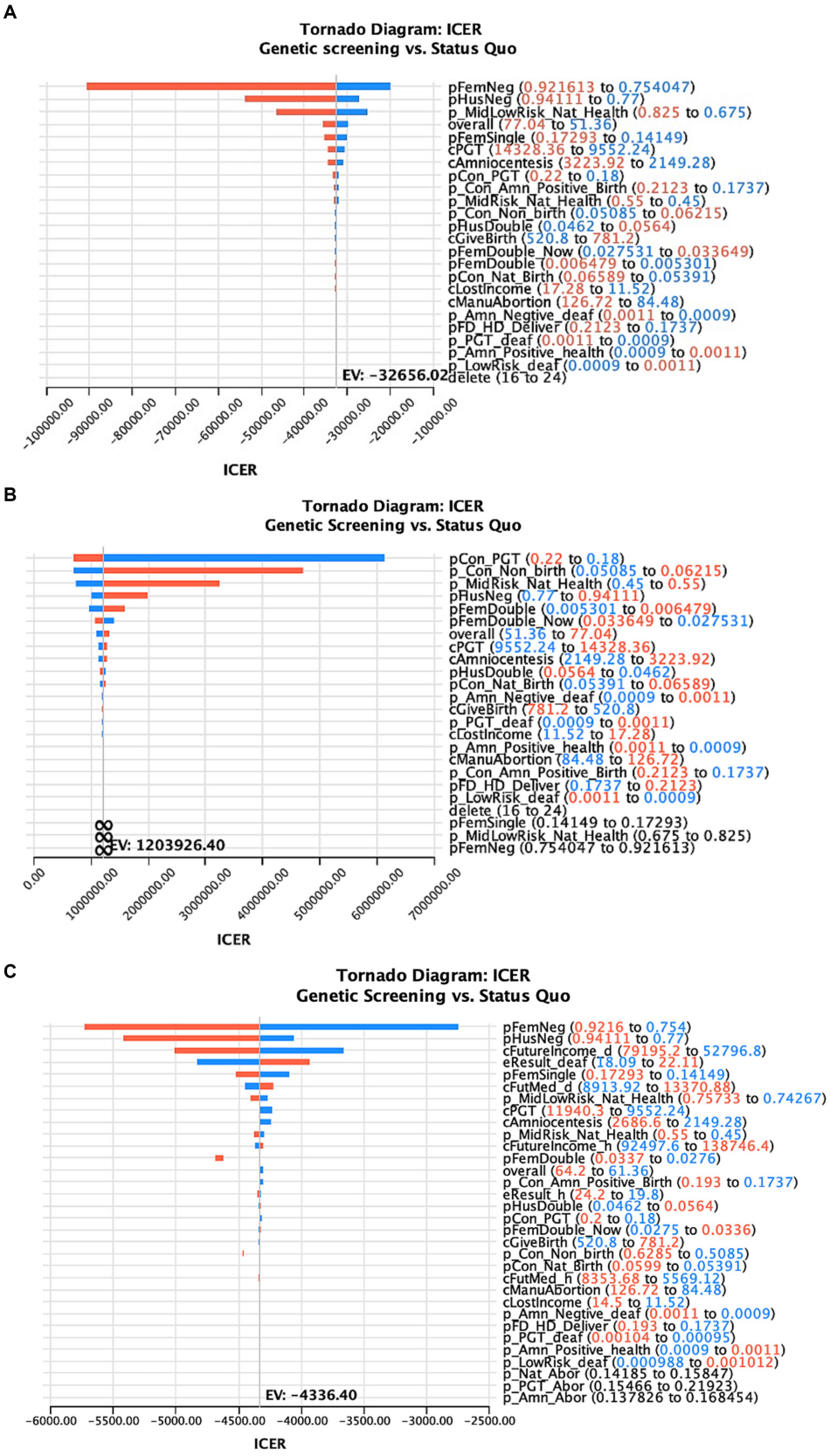


FIGURE 3
One-way sensitivity analysis of the factors affecting the ICER of deafness screening vs. the status quo of three different models (**A–C** represents the results of the model 1–3 in the main text).

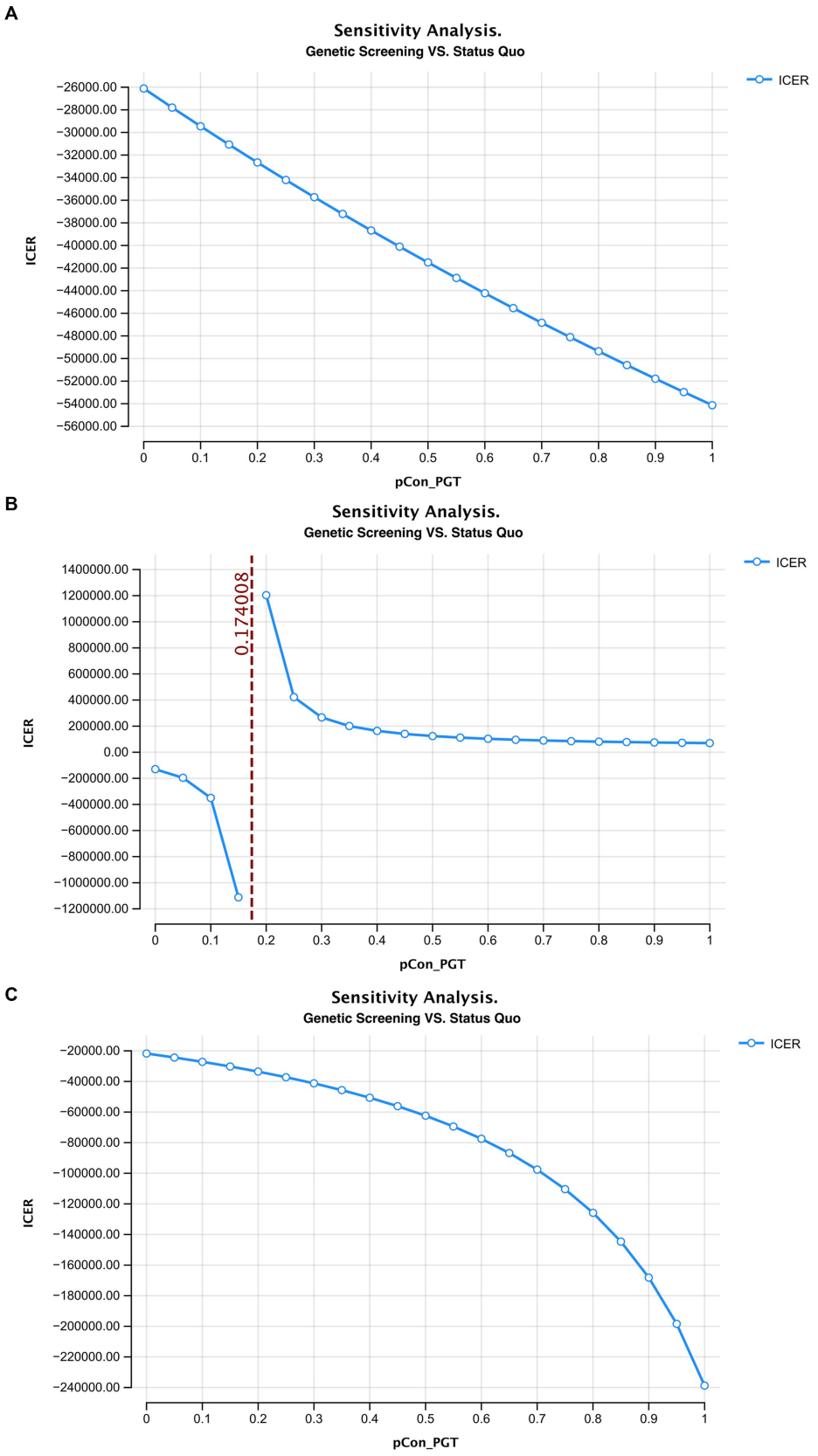


FIGURE 4 (Continued)

10.3389/fpubh.2023.1081339

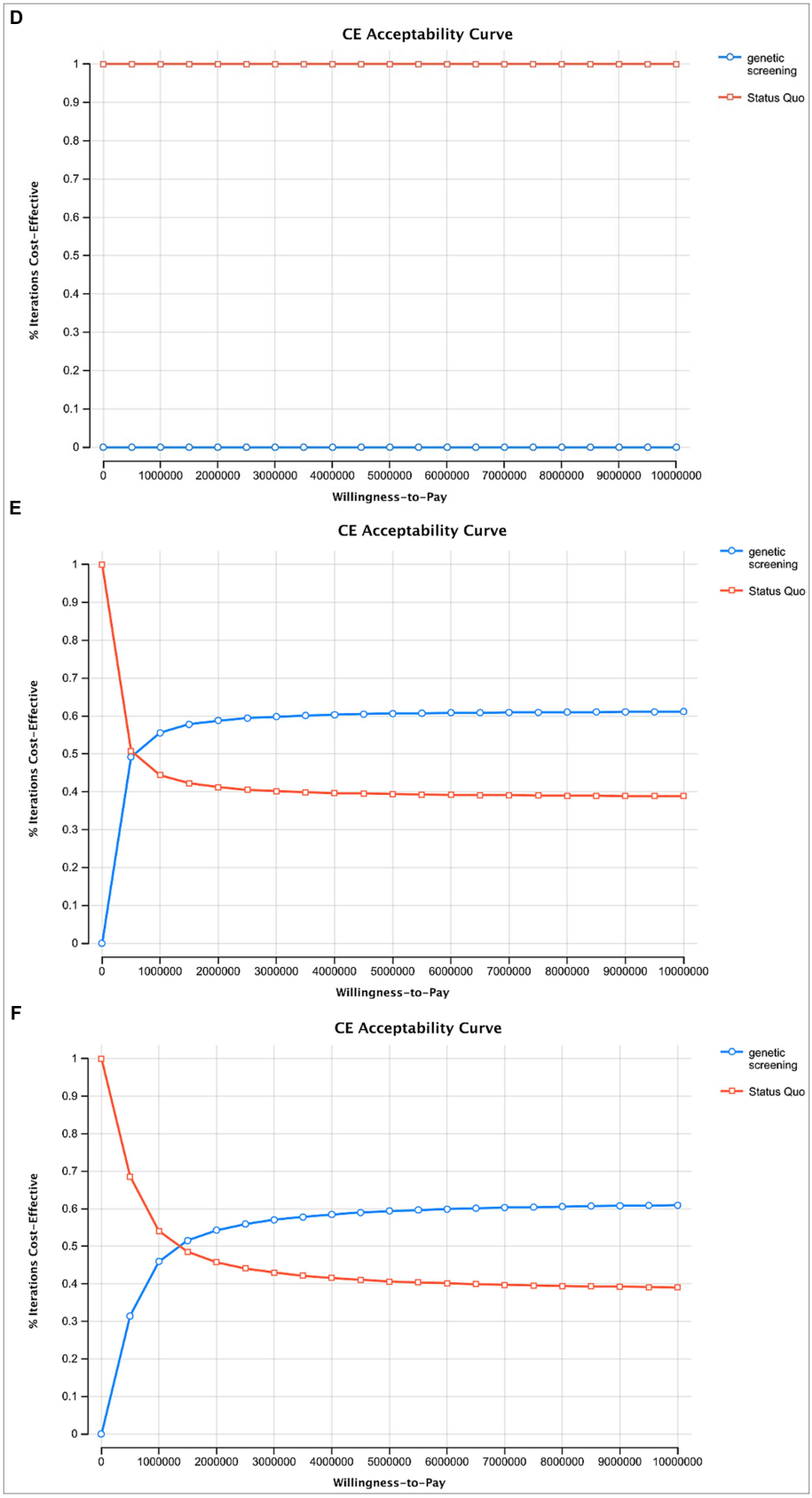


FIGURE 4
The ICER curve and acceptability curve of deafness screening vs. the status quo of the 3 models (**A–C** represents the ICER of the model 1–3 in the main text, **D–F** represents the acceptability curve of the model 1–3 in the main text).

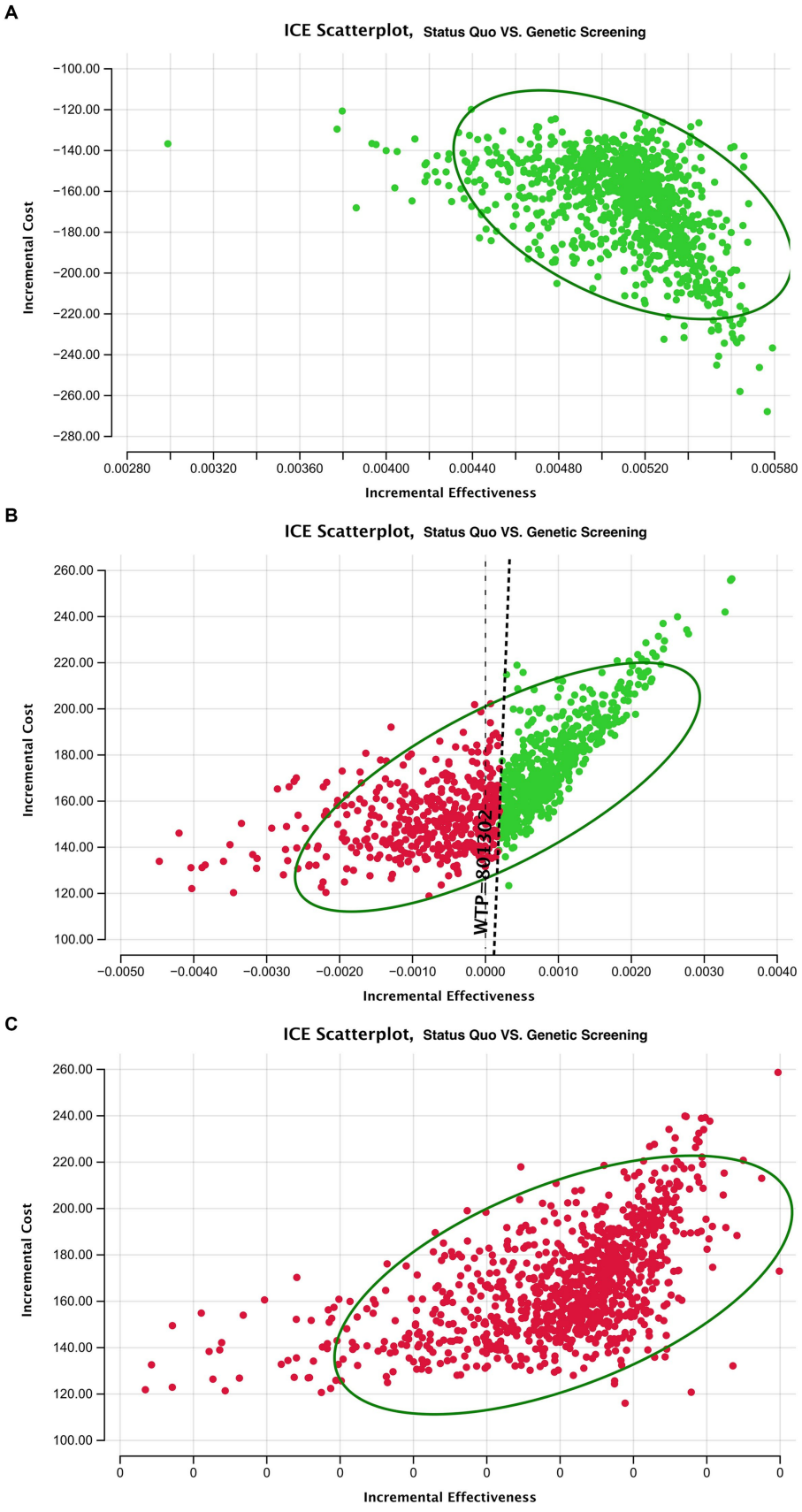


FIGURE 5
ICER scatterplot of sensitivity analysis for deafness screening strategy vs. the status quo of the three models. Circle means 95% of the ICER results after Monte Carol simulation are located in the circle.

There are some limitations of this study. Firstly, more follow-up research is required to collect follow-up data on couples that opt for various interventions after the screening, which may replace the expert consultant and reference citation in the data collection component and can better help inform policy-makers. Secondly, we utilized estimates of future medical expenditure due to deafness; access to accurate estimates would help make our models more accurate. Thirdly, the severity of the hearing loss is likely to worsen over time. A Markov model as a suitable model for periodic change events could be used to explore the effect of genetic screening and its following interventions with more complicated data and research design. Finally, we did not consider the relationship between cost and the families' willingness to undergo screening as it is generally believed that better medical insurance coverage and less personal spending will lead to patients' willingness to receive medical services; this would be valuable for determining medical insurance levels and their effects.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Ninth People's Hospital, Shanghai Jiaotong University School of Medicine. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

YL, ZhiW, ZhaoW, TY and HW made a substantial contribution to the conceptualization and design of the study. YL, YC, and ZhiW collected the data and analysis and interpretation of the data, and drafting of the article. ZhaoW, TY, and YC contributed substantially to the design of the study and drafting and revision of the article. YL, LY, and FC contributed substantially to the interpretation of the data and made critical revisions for the paper. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by National Natural Science Foundation of China (72204156), National Key R&D Program of China (2017YFC1001804), the Shanghai Pujiang Program (2020PJC081), Shanghai Jiao Tong University "Start-up Plan for New Young Teachers" (21X010501094), Soft Science Project of Shanghai Science and Technology Innovation Action Plan (22692192000), Shanghai Municipal Education Commission – Gaofeng Clinical Medicine Grant (20152519), and Shanghai Key Laboratory of Translational Medicine on Ear and Nose Diseases (14DZ2260300).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1081339/full#supplementary-material

References

1. Ministry of Health of the People's Republic of China (2012). China Birth Defects Prevention and Control Report (2012). Available at: http://www.nhc.gov.cn/wsb/pxwfb/201209/55840/files/0af7007b1a68469397531b154d9425f9.pdf?eqid=a4112113000099b70000000464424f01

2. Wroblewska-Seniuk KE, Dabrowski P, Szyfter W, Mazela J. Universal newborn hearing screening: methods and results, obstacles, and benefits. *Pediatr Res*. (2017) 81:415–22. doi: 10.1038/pr.2016.250

3. Ben-Dov T, Brownstein Z, Nageris B, Avraham KB. Innovations in research of hereditary deafness. *Harefuah*. (2020) 159:117–22.

4. Renauld JM, Basch ML. Congenital deafness and recent advances towards restoring hearing loss. *Curr Protoc*. (2021) 1:e76. doi: 10.1002/cpz1.76

5. World Health Organization (2020). Childhood hearing loss: Strategies for prevention and care. Available at: https://iris.who.int/handle/10665/204632 (Accessed November 23, 2023).

6. Smith RJ, Bale JF, White KR. Sensorineural hearing loss in children. *Lancet*. (2005) 365:879–90. doi: 10.1016/S0140-6736(05)71047-3

7. Yang T, Guo L, Wang L, Yu X. Diagnosis, intervention, and prevention of genetic hearing loss. *Adv Exp Med Biol*. (2019) 1130:73–92. doi: 10.1007/978-981-13-6123-4_5

8. Liang J, Mu Y, Li X, Tang W, Wang Y, Liu Z, et al. Relaxation of the one child policy and trends in caesarean section rates and birth outcomes in China between 2012 and 2016: observational study of nearly seven million health facility births. *BMJ*. (2018) 360:k817. doi: 10.1136/bmj.k817

9. Li HT, Hellerstein S, Zhou YB, Liu JM, Blustein J. Trends in cesarean delivery rates in China, 2008-2018. *JAMA*. (2020) 323:89–91. doi: 10.1001/jama.2019.17595

10. National Bureau of Statistics of the People's Republic of China (2018). National Time use Survey Bulletin. 2019. Available at: http://www.stats.gov.cn/tjsj/zxfb/201901/t20190125_1646796.html (Accessed November 23, 2023).

11. National Bureau of Statistics of the People's Republic of China (2022). Statistical Communiqué of the Prc on the 2021 National Economic and social development. http://www.stats.gov.cn/xxgk/sjfb/zxfb2020/202201/t20220117_1826436.html (Accessed November 23, 2023).

12. China Disabled Persons' Federation Research Center for the Development of Disabled Persons & China Disabled Persons' Federation Information Center. National Survey Report on income status of disabled households. *Disabil Stud*. (2019) 2020:75–81.

13. Mcdaid D, Park AL, Chadha S. Estimating the global costs of hearing loss. *Int J Audiol*. (2021) 60:162–70. doi: 10.1080/14992027.2021.1883197

14. Montes F, Peñaranda A, Correa S, Peñaranda D, García JM, Aparicio ML, et al. Cochlear implants versus hearing aids in a middle-income country: costs, productivity, and quality of life. *Otol Neurotol*. (2017) 38:e26–33. doi: 10.1097/MAO.0000000000001393

15. National Bureau of Statistics of the People's Republic of China. *China health statistical yearbook*. Beijing: China Union Medical College Press (2021).

16. Fu X, Cai Y, Hu Y, Liu J, Yang T. Attitudes toward carrier screening and prenatal diagnosis for recessive hereditary deafness among the educated population in urban China. *Am J Med Genet A*. (2016) 170:3180–4. doi: 10.1002/ajmg.a.37932

17. World Health Organization. *World report on hearing. Licence: Cc By-Nc-Sa 3.0 Igo*. Geneva: World Health Organization (2021).

18. Wang JX, Norman RJ, Wilcox AJ. Incidence of spontaneous abortion among pregnancies produced by assisted reproductive technology. *Hum Reprod*. (2004) 19:272–7. doi: 10.1093/humrep/deh078

19. Salomon LJ, Sotiriadis A, Wulff CB, Odibo A, Akolekar R. Risk of miscarriage following amniocentesis or chorionic villus sampling: systematic review of literature and updated meta-analysis. *Ultrasound Obst Gynecol*. (2019) 54:442–51. doi: 10.1002/uog.20353

20. Lee M, Lofgren KT, Thomas A, Lanes A, Goldman R, Ginsburg ES, et al. The cost-effectiveness of preimplantation genetic testing for aneuploidy in the United States: an analysis of cost and birth outcomes from 158,665 in vitro fertilization cycles. *Am J Obstet Gynecol*. (2021) 225:55.e1-.e17. doi: 10.1016/j.ajog.2021.01.021

21. Lipton JH, Zargar M, Warner E, Greenblatt EE, Lee E, Chan KKW, et al. Cost effectiveness of in vitro fertilisation and preimplantation genetic testing to prevent transmission of Brca1/2 mutations. *Hum Reprod*. (2020) 35:434–45. doi: 10.1093/humrep/dez203

22. Huddle MG, Goman AM, Kernizan FC, Foley DM, Price C, Frick KD, et al. The economic impact of adult hearing loss: a systematic review. *JAMA Otolaryngol Head Neck Surg*. (2017) 143:1040–8. doi: 10.1001/jamaoto.2017.1243

23. Kuliev A, Rechitsky S. Preimplantation genetic testing: current challenges and future prospects. *Expert Rev Mol Diagn*. (2017) 17:1071–88. doi: 10.1080/14737159.2017.1394186

24. Ong KIC, Khattignavong P, Keomalaphet S, Iwagami M, Brey P, Kano S, et al. Health-seeking behaviours in a malaria endemic district in Lao People's Democratic Republic: a mixed methods study. *BMJ Open*. (2021) 11:e055350. doi: 10.1136/bmjopen-2021-055350

25. Lattof SR. Health insurance and care-seeking behaviours of female migrants in Accra, Ghana. *Health Policy Plan*. (2018) 33:505–15. doi: 10.1093/heapol/czy012

26. Lieu JEC, Kenna M, Anne S, Davidson L. Hearing loss in children: a review. *JAMA*. (2020) 324:2195–205. doi: 10.1001/jama.2020.17647

27. Brown CS, Emmett SD, Robler SK, Tucci DL. Global hearing loss prevention. *Otolaryngol Clin N Am*. (2018) 51:575–92. doi: 10.1016/j.otc.2018.01.006

28. Bond M, Mealing S, Anderson R, Elston J, Weiner G, Taylor RS, et al. The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: a systematic review and economic model. *Health Technol Assess*. (2009) 13:1–330. doi: 10.3310/hta13440

29. Bittencourt AG, Torre AA, Bento RF, Tsuji RK, Brito Rd . Prelingual deafness: benefits from cochlear implants versus conventional hearing aids. *Int Arch Otorhinolaryngol*. (2012) 16:387–90. doi: 10.7162/S1809-97772012000300014

30. Watkin P, Baldwin M. The longitudinal follow up of a universal neonatal hearing screen: the implications for confirming deafness in childhood. *Int J Audiol*. (2012) 51:519–28. doi: 10.3109/14992027.2012.673237

31. Morton CC, Nance WE. Newborn hearing screening--a silent revolution. *N Engl J Med*. (2006) 354:2151–64. doi: 10.1056/NEJMra050700

32. Fortnum HM, Summerfield AQ, Marshall DH, Davis AC, Bamford JM, Davis A, et al. Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study. *BMJ*. (2001) 323:536–40. doi: 10.1136/bmj.323.7312.536

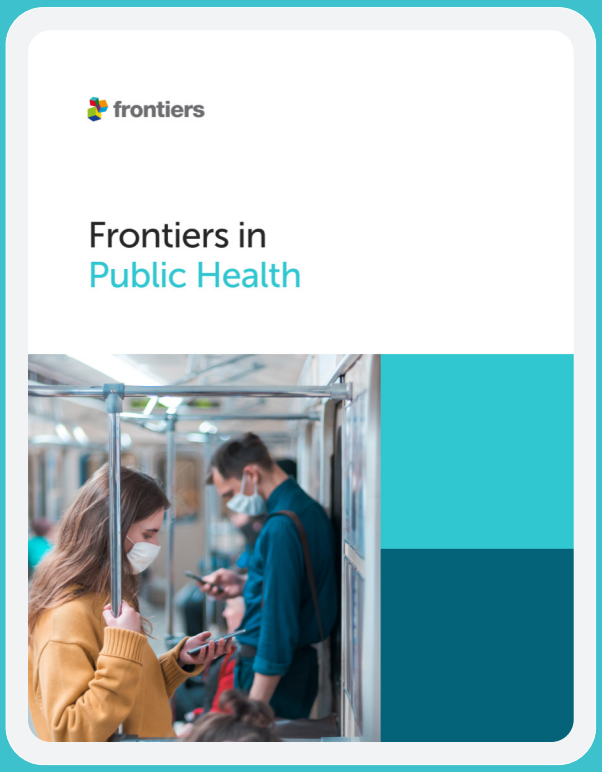
Frontiers in Public Health

Explores and addresses today's fast-moving healthcare challenges

One of the most cited journals in its field, which promotes discussion around inter-sectoral public health challenges spanning health promotion to climate change, transportation, environmental change and even species diversity.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact