Multi-index Evaluation of Health Care System and Its Effect Prediction

Yue Fan 1, a, Yunhan Zhang 2, b, Yun Luo 3, c

1 School of Information Engineering, China University of Geosciences (Beijing), Beijing, China
2 School of Economics and Management, China University of Geosciences (Beijing), Beijing, China
3 School of Geophysics and Information, China University of Geosciences (Beijing), Beijing, China

a 945162824@qq.com, b zhangyunhan@mail.niit.edu.cn, c 1325374050@qq.com

Abstract. Health care systems are crucial to people’s lives, so we established a quantitative model of multi-index health care assessment system to evaluate health care systems of different countries objectively. Firstly, we adopted the entropy weight method to calculate the weight of each indicator and built a linear function to derive the scoring results of the system in each country. Secondly, we used the incomplete induction method for the secondary screening of indicators, and the weight of selected seven indicators are: 0.16 0.11 0.13 0.14 0.19 0.15 0.12. Thirdly, we compared different countries’ health care systems. Compared to the America’s health care system, Japan had a better one, while China had a poorer one, and the scores of them are: 71.23 77.87 38.96. Finally, we proposed two suggestions to U.S. and predicted its improvement effects by BP neural network. Given the ease of implementation, cost and other issues, the latter is considered more appropriate.

Keywords: Health Care System; Entropy Weight Method; Incomplete Induction; BP Neural Network.

1. Background
Currently, researchers have focused on health care system evaluation in two directions: first, from the micro level, to investigate individual health care accessibility [4], i.e., how likely residents are to receive effective treatment after becoming ill; second, from the macro level, to study health care system performance evaluation [12], i.e., the process of assessing the achievement of health care system goals given health resources. Previously, scholars have combined micro and macro research methods to study health care systems and make recommendations, but the AHP analysis used is highly subjective, and the research inferences lack of typicality [2].

Consequently, we will overcome the shortcomings of the above studies and select appropriate indicators to assess health care systems.
2. Assumptions and Justifications
To simplify the model, this paper makes following basic assumption, each of which is properly justified.

- The influence of non-human resistible factors on the fluctuation of the evaluation is not considered. We assume that all countries are safe and free from war, while there will be no small probability events such as floods and earthquakes, therefore removing the influence of chance factors on the model.

- The weight of factors in the model are fixed in the short term and do not fluctuate in different countries. In this way, we assume that the weight of the factors do not change over time in the short term, while guaranteeing that the model can be applied to any country and does not necessarily need to be adjusted according to the country.

- Access to effective therapy is correlated with the wealth status of individuals. Studies show that Americans at all income levels are less healthy than those with incomes higher than their own[1]. Not only is income (earnings and other money acquired each year) associated with better health, but wealth (net worth and assets) affects health as well.

3. Notions
Table 1 is a description

| Symbol | Definition |
|--------|------------|
| z_{ij} | Standardized value of index data of health care evaluation system of various countries |
| x_{ij} | Data on the indicators of the health evaluation system of various countries |
| x_{max} | The maximum value in the data of the indicators of evaluation system of each country |
| x_{min} | The minimum value in the data of the indicators of evaluation system of each country |
| i | The label of any country |
| j | The label of any health care evaluation system indicator |
| p_{ij} | Probability when calculating information entropy |
| E_{ij} | Information entropy |
| n | Number of countries selected in the health care evaluation system |
| k | Number of indicators selected in the health care evaluation system |
| d_{ij} | Information utility value |
| w_{ij} | Weight |
| a_{t} | The score of each country's system calculated by the constructed health care evaluation system |
| S_{i} | The value of each country's system score when magnified by 100 times |
| S_{j} | Score of the health care system of the jth country after deleting the jth indicator |
| P_{j} | The sum of the scores of the difference between the scores of countries before deleting the jth indicator and after deleting the jth indicator |
| l | Number of hidden layer neurons |
| T_{m} | Expected output of BP neural network |
| O_{m} | Calculation output of BP neural network |

4. Model 1: Quantification of Health Assessment Model Based on Entropy Weight Method

4.1. Data Selection
By consulting the literature [2], we selected 13 indicators from the World Health Statistics 2020[3] and the global medical access and medical quality list published by The Lancet to evaluate the health care system [4]. The 13 indicators cover a wide range, including official coefficients, high-risk diseases, maternal mortality, living environment, medical expenditures, medical conditions, and personal habits. The health care system can be evaluated at a subjective level.

Indicators include age-standardized prevalence of tobacco use among persons 15 years and older(PT), total alcohol per capita consumption (litres of pure alcohol)(TAC), maternal mortality ratio(per 100000 live births)(MMR), neonatal mortality rate (per 1000 live births)(NMR), population with household expenditures on health >10% of total household expenditure or income(PET), HAQ index 2016(HAQ),
NM skin cancer (SCC) cure rate(SCC), Leukaemia cure rate(LC), healthy life expectancy at birth(years)(HLE), density of medical doctors (per 10000 population)(DM), density of nursing and midwifery personnel (per 10000 population)(DN), density of dentists (per 10 000 population)(DD) and domestic general government health expenditure (GGHE-D) as percentage of general government expenditure(DGP).

The 13 selected indicators were divided into positive and negative indicators according to the impact of each indicator on the health care system assessment model. As shown in Figure 1

We chose 16 developed and developing countries from 6 continents, including Australia, Canada, China, Congo, Germany, Haiti, India, Italy, Japan, Mongolia, Peru, Russia, Rwanda, Singapore, Thailand and USA.

4.2. Application of Entropy Weight Method in Establishment

According to the explanation of the basic principle of information theory, information is a measure of the degree of system order, and entropy is a measure of the degree of systemic disorder [5]. So, the information entropy tool can be used to calculate the weight of each indicator, providing a basis for the comprehensive evaluation of multiple indicators. We adopt entropy weight method to obtain different weight of different indicators on the results, form scoring rules, quantify the contribution of each indicator, and then we can analyze the contribution of each index. Based on this principle, entropy weight method can objectively evaluate health care systems. The entropy weight method comprises the following steps.

4.2.1. Standardize the data of each indicator. Standardize the data of each indicator. Here, Min-Max standardization is used to process the corresponding data of each indicator to ensure the non-negativity of the data.

![Figure 1. Positive indicators and transparency indicators](image-url)
Among them, $z_{ij}$ is the standardized variable. $x_{max}$ and $x_{min}$ are the maximum and minimum values of each indicator respectively.

4.2.2. Calculate the information entropy of each index. We calculate the weight of the $i$th country under the $j$th health care indicator. Then view it as the probability in calculating the information entropy $p_{ij}$. Here, $n$ is the number of countries and $k$ is the number of indicators.

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij}}$$

Compute the information entropy $E_j$ of the $j$th health care indicator, and calculate the corresponding information utility value $d_j$. According to the definition of information entropy in information theory, the information entropy of a set of data is

$$E_j = -\ln \left( \frac{1}{n} \sum_{j=1}^{n} p_{ij} \ln p_{ij} \right)$$

The information utility value is

$$d_j = 1 - E_j$$

If

$$p_{ij} = 0$$

then

$$\lim_{p_{ij} \to 0} p_{ij} \ln p_{ij} = 0$$

4.2.3. Determine the weight of each indicator

The weight of each indicator $w_j$ can be obtained by normalization.

$$w_j = \frac{d_j}{\sum_{j=1}^{k} d_j}$$

4.3. The Result

We get the weights of these 13 indicators as the table 2 below shows.

| Indicator | Weight |
|-----------|--------|
| PT        | 0.052  |
| TAC       | 0.110  |
| MMR       | 0.032  |
| NMR       | 0.047  |
| PET       | 0.038  |
| HAQ       | 0.075  |
| SCC       | 0.102  |
| LC        | 0.085  |
| HLE       | 0.061  |
| DM        | 0.091  |
| DN        | 0.124  |
| DD        | 0.101  |
| DGP       | 0.083  |

First, we can see that DN has the highest weight, which is close to our expectations. This can be explained as the higher the density of nursing and midwifery personnel (per 10000 population), the higher the probability of people getting medical care, which means the better the health care system of
a country. In addition, the weighting ratio of DM and DD is about 10%, which further confirms this view.

Second, the weight of domestic general government health expenditure (GGHE-D) as percentage of general government expenditure (DGP) reached 8%, which may indicate that the capital investment of a country in the health care system can increase the salaries of medical staff and update medical equipment, thus ensuring people’s physical and mental health. Since DGP is an important factor in our society, the result of weight is more reliable.

Finally, indicators such as MMR, NMR, and PET have lower weights, which means that they have less impact on the health care system.

5. Model 2: Quantitative Evaluation Model of Index Selection Based on Incomplete Induction

5.1. Design of the Model
In model 2, we now use the incomplete induction model to select the most vital metrics.

In this model, the evaluation of health care system is an induction where the set of instances is not exhaustive. Under the framework that is considered to be correct, incomplete induction can help us extract the main indicators that affect the quality of the health care system.

First, we choose for our analysis 16 countries, from different regions of different continents, from different levels of development, and with different health care systems. In other words, they are representative in the whole world: Australia, Canada, China, Congo, Germany, Haiti, India, Italy, Japan, Mongolia, Peru, Russia, Rwanda, Singapore, Thailand and USA.

Then, indicators are divided into 7 groups, as shown in Figure 2 below.

![Incomplete induction model to select the most important indicators](image)

**Figure 2.** Incomplete induction model to select the most important indicators

5.2. Data processing
For the ith indicator of the ith country, the score of this indicator is obtained after standardization according to the existing data as $z_{ij}$. Set the standardized score of the ith country as $a_i$, then

$$a_i = \sum_{j=1}^{k} w_j z_{ij}$$
Amplify the score by 100 times to distribute it in the interval of 0-100 to make the data easy to observe. Suppose the score of the ith country is \( S_i \), then

\[
S_i = 100a_i
\]

The health care system scores of 16 countries are shown in Figure 3.

![Figure 3](image)

**Figure 3.** The health care system scores of 16 countries

Delete the jth metric in turn, and calculate the score \( S_0 \) of 16 countries with the same calculation method using other 12 indicators.

### 5.3. Filtering main indicators

Let

\[
P_j = \sum_{i=1}^{n} (S_i - S_j)^2
\]

Choose the metrics associated with the three (or more) largest \( P_j \).

We arrange the P value from large to small, and select 6 indicators (total alcohol per capita consumption (litres of pure alcohol), HAQ index 2016, Leukaemia cure rate, density of medical doctors (per 10000 population), density of nursing and midwifery personnel (per 10000 population) and density of dentists (per 10 000 population)) of the four categories with larger \( P_j \) values, which can effectively evaluate the health care system of each country.

The incomplete induction method requires that each category must contain at least one indicator, if not, adjustment should be made according to the actual situation. In order to ensure the integrity of the model, on the basis of the above 6 indicators, we added domestic general government health expenditure (DGGHE-D) as percentage of general government expenditure (DGP), whose \( P_j \) value is 346.71. In this way, it can be guaranteed that at least one indicator is contained in the medical expenses category.

Finally, we obtained 7 indicators with the following weights, as shown in Table 3 below.

| TAC | HAQ | LC | DM | DN | DD | DGP |
|-----|-----|----|----|----|----|-----|
| 0.16| 0.11| 0.13| 0.14| 0.19| 0.15| 0.12|
The calculation formula of the total attributional indicator (TAI) obtained by incomplete induction is shown as follows.

\[ TAI = 0.16 \times AC + 0.11 \times HAQ + 0.13 \times LC + 0.14 \times DM + 0.19 \times DN + 0.15 \times DD + 0.12 \times DGP \]

Based on the final 7 indicators that were chosen, the adjusted scores of 16 countries are shown below. The adjusted score is shown in Figure 4.

5.4. Comparison

5.4.1. The United States vs. Japan. According to our model, the U.S. health care system has a rating of 71.23 while Japan has a rating of 77.87. The U.S. health care system is also rated less favorably than Japan in the WHO 2000 assessment.

In Japan, services are provided either through regional/national public hospitals or through private hospitals/clinics, and patients have universal access to any facility, though hospitals tend to charge more to those patients without a referral. As above, costs in Japan tend to be quite low compared to those in other developed countries, but utilization rates are much higher. Japanese patients favor medical technology such as CT scans and MRIs, and they receive MRIs at a per capita rate twice higher than Americans [7]. In most cases, CT scans, MRIs and many other tests do not require waiting periods. Japan has about three times as many hospitals per capita as the US and, on average, Japanese people visit the hospital more than four times as often as the average American [8].

Therefore, based on the above results, we conclude that Japan has a better healthcare system.

5.4.2. The United States vs. China. In this section, we choose the United States with a score of 71.23 to compare with China with a score of 38.96 and use the latest available data to prove our inference that there is a gap between the health care systems of China and the United States.

Some researchers found various differences between China and the USA with respect to medical education, hospital management, medical insurance, and communication between patients and doctors. China still faces challenges and gaps in its health education and hospital management system as compared to the USA [9].
6. Model 3: Prediction of American Medical Health System Improvement Based on BP Neural Network

We choose the health care system in the United States to build a predictive model and propose improvements. The BP neural network is a multi-layer pre-feedback neural network. The transmission of neurons is an S-shaped function, and the output is a continuous amount between 0 and 1. It can realize any nonlinear mapping from input to output.

6.1. Data Processing

Based on the difficulty of data collection and the degree of influence of each indicator on the health care assessment system, we selected six indicators, which are TAC, LC, DM, DN, DD, DGP. We choose the premmnx function of Matlab to normalize the data for the jth metric to zmj.

6.2. Model Establishment

A BP neural network consists of an input layer, a hidden layer, and an output layer, and the hidden layer can have one or more layers. It has been shown that a neural network with one hidden layer can approximate a nonlinear function with arbitrary accuracy as long as there are enough hidden nodes [11], therefore, a three-layer multi-input single-output BP neural network containing one hidden layer is chosen to build the prediction model.

According to the number of indicators, we choose six neurons as the input layer in the model construction process. We take HAQ index as the target of the network, therefore there is only one neuron in the output layer. As shown in Figure 5

![Figure 5. Output layer](image)

In the process of network design, to determine the number of hidden layer neurons, there is an empirical formula as follows.

\[ l = \sqrt{n + m + a} \]

Here, n is the number of neurons in the input layer, m is the number of neurons in the output layer, and a is a constant in the interval 1 to 10.

According to the empirical formula, we can calculate that the number of neurons in the hidden layer belongs to the interval 4 to 13. Combining multiple experiments, we have selected 9 neurons as the hidden layer.

As a result, we derived a 6*9*1 three-layer BP network model with the structure shown in Figure 5. According to the design principles of the BP neural network, we choose S-type logarithmic function \( f(x) = \frac{1}{1+e^{-x}} \) as the activation function of hidden layer and output layer. Their normalized range is
[0, 1]. By back propagation error function \( E = \frac{1}{2} \sum_{m} (T_m + Q_m)^2 \) the weight and threshold of the network are constantly adjusted to make the error function \( E \) minimized.

6.3. Results and Suggestions

According to the data we found for the six indicators of TAC, LC, DM, DN, DD, and DGP from 2010 to 2018, we predict the HAQ of the United States from 2010 to 2018. The results are as follows Table 4.

Table 4. Prediction of HAQ

| Year | Predicted HAQ |
|------|---------------|
| 2010 | 67.40         |
| 2011 | 68.50         |
| 2012 | 69.00         |
| 2013 | 73.80         |
| 2014 | 78.00         |
| 2015 | 81.00         |
| 2016 | 88.70         |
| 2017 | 88.90         |
| 2018 | 89.00         |

It can be seen that the HAQ index in the United States is increasing year by year, and the health care system is becoming more and more perfect.

We set a 10% reduction in total alcohol per capita consumption (litres of pure alcohol) (TAC) and a 10% increase in domestic general government health expenditure (GGHE-D) as percentage of general government expenditure (DGP) as recommendations.

By changing the input value, the BP neural network completes the learning after reaching the expected error through repeated learning. We come to our prediction results as follows. As shown in Figure 6.

7. Conclusion

In order to evaluate the country’s health care system, this article initially screened thirteen indicators, and then selected seven indicators through Model 2 for comprehensive evaluation of different countries.

Japan, which generally believes that its health care system is better than the United States, scores higher in our model than that of the United States; China, which generally believes that its health care system is slightly worse than the United States, scores lower in our model than that of the United States.
Taking into account the practical factors such as the difficulty of improving the medical system, the promotion intensity, and the cost, the article proposes suggestions for TAC and DGP in the United States. Through the prediction of Model 3, it is found that 10 Percent Reduction in TAC and 10 Percent Increase in DGP It can effectively improve the score of the US Healthcare Access and Quality Index.

References
[1] Braveman PA, Cubbin C, Egerter S, et al. Socioeconomic Disparities in Health in the United States: What the Patterns Tell Us[J]. American Journal of Public Health, 2010, 100 Suppl 1(supplement 1): S186-96.
[2] Suo Yiming, Gu Qi, Li Li. Comparative analysis of health care systems in China, the United States and Germany and policy recommendations for China—a mathematical modeling perspective[J]. Modern Preventive Medicine, 2009(12):2299-2302.
[3] World Health Organization. World Health Statistics 2020[EB/OL].
url {https://www.who.int/data/gho/publications/world-health-statistics}
[4] Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016[J]. Lancet, 2018, 391(10136):2236-2271.
[5] Cheng Qiyue. Structural entropy weight method for evaluating index weights [J]. Systems engineering theory and practice, 2010, 30(7):1225-1228.
[6] Wikipedia. World Health Organization ranking of health systems in 2000[EB/OL].
url {https://en.wikipedia.org/wiki/World_Health_Organization_ranking_of_health_systems_in_2000}
[7] Frontline. Sick around the world[EB/OL].
url {https://www.pbs.org/wgbh/pages/frontline/sickaroundtheworld/}, April 15, 2008.
[8] Harden, Blaine. Health Care in Japan: Low-Cost, for Now Aging Population Could Strain System[EB/OL]. The Washington Post, January 28, 2010.
[9] Medical education and health care in China and the USA: a follow-up study of a 4 year cardiovascular visiting programme between Peking University and the University of Minnesota
[10] Yan Zhaoqiang. Application of BP Neural Network in Mathematical Modeling
url {http://www.doc88.com/p-9475015275086.html}, 2018-12-04.
[11] Data statistical analysis and mining. Detailed BP neural network prediction algorithm and implementation process example[EB/OL].
url {https://blog.csdn.net/weixin_45620071/article/details/100811845}, 2019-09-13.
[12] Tan Peng, Dai Tao, Fu Hongpeng, et al. The characteristics and enlightenment of the performance evaluation framework of the international health system[J]. Chinese Health Policy Research, 2018, 12(004):6-12.