Submitted: 25 January, 2021 Accepted: 09 February, 2021 Published: 30 September, 2021 DOI:10.31083/jomh.2021.024

JOMH Journal of Men's Health

Original Research

Investigate the relationship between men's health and medical publication

Na Ran[1,*]

[1]Library, University of Science and Technology Beijing, 30 Xueyuan Road, Haidian District, 100083 Beijing, China

***Correspondence:** ranna@ustb.edu.cn (Na Ran)

Abstract

Background and objectives: The role of medical papers in different disciplines in men's health has not been well-clarified. The aim was to study qualitatively and quantitatively the impacts of medical papers in different disciplines of medical science on men's health.

Materials and methods: The data have been analyzed by combining correlation analysis, principal component analysis, and multivariate regression methods. This combination method can clean input data by using its correlation with the target data. Thus, the weight of an input parameter by using the multivariate regression method can be obtained. Thus, the proposed method is used to perform personalized optimization.

Results: 69.3% of the variation ($P < 0.01$) in the male mortality rate of China can be explained by the first principal component. 75.8% of the variation ($P < 0.01$) in the male life expectancy at birth in China can be explained by the first and the second principal component.

Conclusion: The published medical papers in different disciplines, education, income, and GDP have equally important effects on men's health.

Keywords

Men's health; Medical publishing; Principal component analysis; Regression methods; Correlation coefficient

1. Introduction

Sustainable development might be described by the complex interrelationships between natural and social spheres [1]. It needs organizations to make proper managerial decisions and realizing actions and behaviors by getting the necessary knowledge and using it with maximum effectiveness [2]. 'Dissemination', 'utilization', 'evidence into practice', and 'knowledge transfer' are principles of knowledge translation [3, 4]. Effective knowledge translation in indigenous contexts needs understanding local indigenous processes of knowledge creation, dissemination, and utilization [5]. What is and not considered being knowledge translation is the most important [6]. Knowledge translation in some circumstances had been found as effective as complex and multifaceted ones [7]. Definitions of knowledge translation are unclear [8, 9]. This leads that information retrieval is difficult [10]. Public knowledge organizations support governments, professionals, and businesses by collecting, translating, and disseminating knowledge [11]. Knowledge translation in the health field strategies involves a series of public or community prevention alliance strategies in health and well-being disciplines [12], drug misuse prevention services for young people [13], promotion of healthy weight [14], immunization and cancer prevention [15]. The research utilization theory of knowledge translation shows that knowledge is a set of constantly changing understanding formed by those who produce and use research [16]. It inferred that potential users are more likely to do so when there is a clear need or motivation [17]. This is similar to the theory of innovation diffusion [18]. Potential adopters of innovation can be divided into innovators, early adopters, early majority, and so on [19]. Many theories have different goals, from providing information for individuals or large

audiences to achieve behavior change through education or skill acquisition [20]. The success, challenges and lessons of using social media in health research had been studied [21].

The health benefits of physical activities need the necessary knowledge, skills, and so on [22]. The capacity of Inuit to be on the ice safely can be influenced by the change in the sea ice travel knowledge. These changes are transforming the sea ice for Inuit from a place, which is an important source of health [23]. Understanding the colorectal cancer screening tests can increase the related confidence and benefit cognition [24]. It is important to increase knowledge and awareness by training health professionals to communicate and provide targeted pre-pregnancy care [25]. Health promotion is not only related to disease prevention but also needs to be invested in health education aimed at knowledge sharing and knowledge development [26]. The delays because of lacking the knowledge to diagnose or suspect breast cancer can lead to worse results [27]. Compared to their less-educated peers, those with higher education had lower mortality, experienced less harmful diseases, and felt healthier [28]. Education improves a person's health knowledge, enables the educated to choose a more effective investment of health products, and leads to improve health results [29].

Men who were screened for prostate-specific antigen had a lower risk of prostate cancer. Knowledge of the disease is an effective cause of men taking part in the screening program because research shows that men with higher knowledge levels have higher screening tendencies. Besides, studies have shown that mass media such as radio and journals are their sources of prostate cancer screening knowledge [30]. More importantly, the Internet might have an impact on the population's medical knowledge. Before getting help from a doctor, people still prefer to use the Internet search to get health information [31]. Medical knowledge is needed for health promotion [32]. Health knowledge is necessary to change behavior [33]. A significant relationship of health literacy with Type 2 Diabetes Mellitus dependent on both knowledge level and education level had been found [34].

It is often difficult to measure the effects of knowledge creation and diffusion on society [35] and it is still unclear about the impact of health knowledge on health status [36]. One can note that most of the published medical papers are open-access articles. At the same time, the Internet has been widely used, and the Internet search is a priority to people seeking health information before getting help from a doctor [31]. In other words, the roles of medical knowledge on men's health have not been well-established. Further qualitative and quantitative investigation on medical publication on men's health is needed to improve men's health. This article aims to do a pioneer investigation into the effects of knowledge translation from the research results published in the scholarly medical journals on men's health.

2. Materials and methods

To check theories and hypotheses, it is important to use data to test those theories and hypotheses. To test hypotheses, two variables (a proposed cause, and a proposed result) need to be measured. After the research data have been collected, what the general trends in the data are via graphic data and fitting the data by using models should be analyzed. Many researchers had used mathematical methods and statistical models to study issues in the social sciences. For example, regression analysis, correlation analysis, and component analysis had been widely used in social science research [37–43]. For developing a personalized optimization method based on principal component analysis, the input data will be cleaned according to its correlation with the target data. The correlation coefficient can be calculated by

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \quad (1)$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $\bar{y} = \sum_{i=1}^{n} y_i$. After the input data have been cleaned oriented to a target, we will perform the principal component analysis.

The curse of dimensionality (the number of input variables) is that when the volume of the space increases fast with the number of input variables increasing, and thus the available data are sparse. Dimensionality reduction is a method to reduce the data from a high-dimensional space into a low-dimensional space. Thus, the number of input variables can be reduced by finding a smaller set of new variables. Many techniques can be used for dimensionality reduction, for example, wavelet transforms, principal components analysis, and attribute subset selection [44]. The principal component analysis is just one technique for dimensionality reducing the of such data sets by finding a smaller set of new variables through a linear transformation. In other words, it is the process of computing the principal components and using them to perform a change of basis of the data, and it can increase interpretability [45]. The principal component analysis uses the largest variance as the main feature, and "dissociates" the data in each orthogonal direction. That is, makes them not correlate different orthogonal directions. After completing the principal component analysis on the data, the correlated features of the data will be removed. It means that all the principal components are independent of each other. Because no correlation between all the principal components, the calculation can be sped up by getting rid of correlated variables that don't contribute to any decision-making. After the data have been implemented by principal component analysis, the original features transform from their linear combination named as principal components. These values of the linear combination are not as legible and interpretable. Also, when principal components are used, some information on the original features might lose. In a word, principal component analysis is a parameterless technology. That is to say, in the face of the same data, if cleaning data is not considered, the results will be the same. Without the intervention of subjective parameters, the principal component analysis is easy to implement in

JOMH
Journal of Men's Health

general, but it can not be personalized optimization [46]. The principal component analysis, a mathematical method, can help us find relationships between two variables sets (a cause variable set and a result variable set) that have been collected for an issue in societies. For example, such a method was used to get socioeconomic impact [47]; was used to investigate among adolescents' academic performance, drug use, sleep quality, and risk of anxiety and depression [48]; was also used for analyzing the performance of semiconductor devices [49]; was used to investigate the relationships between some pre- and post-slaughter traits of broilers [50]; was also used for early disease detection [51]; was also used to predict ozone concentrations [52]. A set of variables with a linear association for two or more explanatory variables can be solved by using the principal component analysis. Thus, Eigenvalues, its variance, the cumulative variance, the loading, and the component score in the principal component analysis can be obtained. The loading indicates the importance of the original variables in the formation of new variables. The component score represents the weighting of new variables in the principal component.

A multiple linear regression between the principal components and the target data will be performed by

$$y = a + \sum_{i=1}^{n} b_i x_i + \varepsilon \tag{2}$$

where b_i, a, and x_i are the regression coefficient for the i-th independent, constant, the i-th independent variables, respectively.

3. Results

Fig. 1 sketches that male life expectancy at birth, current health expenditure, nurses and midwives, school enrollment at tertiary education, school enrollment at tertiary education for female, adjusted savings: education expenditure, GDP per capita, and adjusted net national income per capita in China are found to fluctuate but has a growth trend. The adult male mortality rate is found to decrease with time. Out-of-pocket expenditure is found to swing but has a decreasing trend. This trend infers that education. GDP could have a positive contribution to life expectancy at birth and a negative contribution to the mortality rate. It agrees that education can exert a high impact on health promotion [25, 26, 28, 29, 53]. In this article, either the life expectancy at birth or the mortality rate is used to measure the health and well-being of China.

Fig. 2 shows how documents in health professions, documents in immunology and microbiology, documents in medicine, documents in neuroscience, documents in nursing, documents in pharmacology toxicology pharmaceutics, documents in psychology, documents in veterinary published by China change with time since 1996. All documents in medical journals published by China fluctuate with a growth trend. Such trends imply that medical papers in various disciplines could have a positive

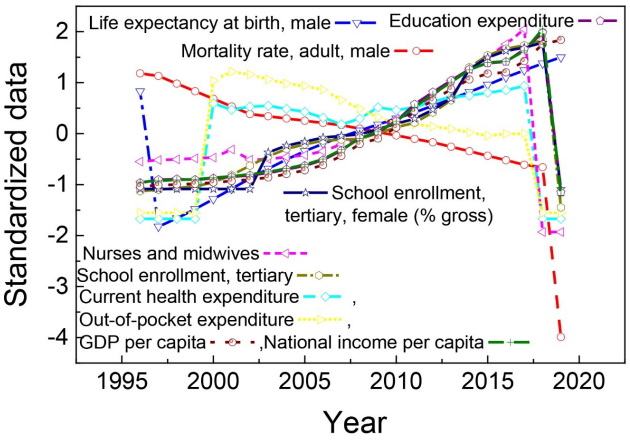

FIG. 1. Annual change in the target data related mens' health and part of input data. Standardized mortality rate (adult, male, per 1,000 male adults), life expectancy at birth (male, years), current health expenditure (% of GDP), nurses and midwives (per 1,000 people), out-of-pocket expenditure (% of current health expenditure), school enrollment (tertiary, % gross), school enrollment (tertiary, female, % gross), adjusted savings: education expenditure (current US$), GDP per capita (current US$), and adjusted net national income per capita (current US$) in China as a function of time. The raw data come from the website of the World Bank (https://data.worldbank.org).

contribution to male life expectancy at birth and a negative contribution to the male mortality rate. It agrees that knowledge is necessary for health promotion [22, 27].

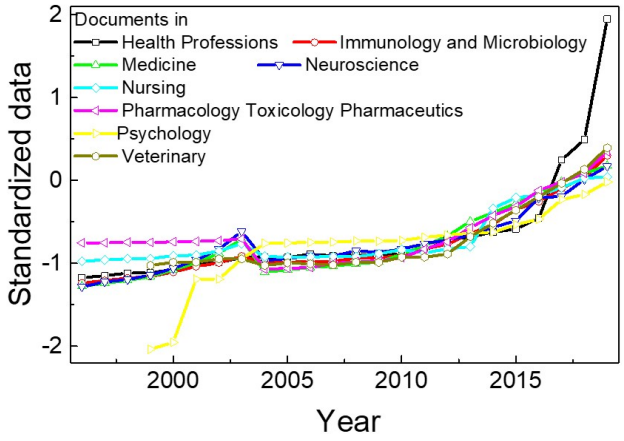

FIG. 2. Annual change in medical publication. Standardized documents in health professions, documents in immunology and microbiology, documents in medicine, documents in neuroscience, documents in nursing, documents in pharmacology toxicology pharmaceutics, documents in psychology, documents in veterinary published by China as a function of time. The raw data come from the website of the website of SCImago (https://www.scimagojr.com).

The Pearson correlation coefficient r is shown in Table 1. The male adult mortality rate (per 1000 male adults) has a significant linear relationship to GDP per capita, documents in health professions, documents in immunology and microbiology, documents in medicine, documents in neuroscience, documents in nursing, documents in pharmacology toxicology pharmaceutics, documents in psychol-

TABLE 1. The Pearson correlation coefficient (*r*). The raw data come from the website of the World Bank (https://data.worldbank.org) and the website of SCImago (https://www.scimagojr.com).

Variables	Mortality rate, adult, male (per 1,000 male adults)			Life expectancy at birth, male (years)		
	r	Sig.	critical value	*r*	Sig.	critical value
Current health expenditure (% of GDP)	0.014	0.948	0.050	0.187	0.380	0.047
Nurses and midwives (per 1,000 people)	0.083	0.699	0.047	0.308	0.144	0.044
Out-of-pocket expenditure (% of current health expenditure)	0.165	0.441	0.044	-0.136	0.525	0.050
School enrollment, tertiary (% gross)	-0.213	0.317	0.041	0.664	0.000	0.041
School enrollment, tertiary, female (% gross)	-0.286	0.176	0.031	0.721	0.000	0.031
Adjusted savings: education expenditure (current US$)	-0.257	0.225	0.034	0.699	0.000	0.034
GDP per capita (current US$)	-0.776	0.000	0.025	0.866	0.000	0.009
Adjusted net national income per capita (current US$)	-0.247	0.245	0.038	0.696	0.000	0.038
Documents in Health Professions	-0.728	0.000	0.028	0.776	0.000	0.028
Documents in Immunology and Microbiology	-0.801	0.000	0.019	0.865	0.000	0.013
Documents in Medicine	-0.834	0.000	0.016	0.869	0.000	0.003
Documents in Neuroscience	-0.834	0.000	0.013	0.844	0.000	0.022
Documents in Nursing	-0.843	0.000	0.006	0.850	0.000	0.019
Documents in Pharmacology Toxicology Pharmaceutics	-0.835	0.000	0.009	0.868	0.000	0.006
Documents in Psychology	-0.885	0.000	0.003	0.669	0.000	0.025
Documents in Veterinary	-0.786	0.000	0.022	0.787	0.000	0.016

ogy, and documents in veterinary because $P < 0.05$. The male adult mortality rate is not significantly linear related to current health expenditure, nurses and midwives, out-of-pocket expenditure, school enrollment at tertiary education, school enrollment at tertiary education for females, adjusted savings: education expenditure, and adjusted net national income per capita because $P > 0.05$. Especially, the GDP per capita, documents in health professions, documents in immunology and microbiology, documents in medicine, documents in neuroscience, documents in nursing, documents in pharmacology toxicology pharmaceutics, documents in psychology, and documents in veterinary have a strong negative correlation to the male adult mortality rate. The observed significant negative correlations display that developing GDP and medical research in various can decrease the male mortality rate. It supports a negative association between economic growth and mortality rates [54]. Table 1 also gives that male life expectancy at birth has a significant linear relationship to school enrollment at tertiary education, school enrollment at tertiary education for female, adjusted savings: education expenditure, GDP per capita, adjusted net national income per capita, documents in health professions, documents in immunology and microbiology, documents in medicine, documents in neuroscience, documents in nursing, documents in pharmacology toxicology pharmaceutics, documents in psychology, documents in veterinary because $P < 0.05$. The male life expectancy at birth is not significant linear related to current health expenditure, nurses and midwives, out-of-pocket expenditure because $P > 0.05$. The observed significant negative correlations demonstrate that developing GDP and medical research in various can decrease the male mortality rate. It supports a positive association for economic growth and life expectancy that had been well-established [55]. One can note that increasing knowledge and developing knowledge can improve public health had been reported [25, 26]. It infers that increasing medical knowledge and developing medical knowledge can lower the mortality rate and increase the life expectancy at birth. There are very strong positive linear relationships between the mortality rate and the article numbers in different medical disciplines, which mean that the mortality rate decreases with the increasing documents. There are very positive linear relationships between the life expectancy at birth and the article numbers in different medical disciplines, which denote that the life expectancy at birth increases with the increasing documents. All these agree well with the former conclusion that increasing knowledge and developing knowledge can improve public health has been reported [25, 26]. Benjamini-Hochberg critical values given in Table 1 have been obtained by controlling the false discovery rate for 0.05 level (Q = 0.05) can be calculated. The largest P-values that are smaller than the critical value of male life expectancy at birth and male mortality rate are 0.00000834 (it corresponds to GDP) and 0.00039798 (School enrollment, tertiary), respectively. In Table 1, these values have been written as 0.000, which should be read as "less than 0.001". At the same time, for male life expectancy at birth and male mortality rate, the P-values of GDP per capita and School enrollment (tertiary) are the largest in those P-values being 0.000 in Table 1. Each P-value smaller than such a largest P-value that is smaller than the critical value is significant.

Collected multidimensional data shown in Fig. 1 and Fig. 2 are likely to be related, in other words, collected multidimensional data have commonness. To see the utility of various kinds of collected multidimensional information data, it is obvious that this commonness should be removed. However, we can not roughly remove the relevant information data (those data are significant), because reducing relevant information data will inevitably lose a lot of important information, which leads to the existence of no one in the effectiveness and reliability of the target information. In addition, how to quickly process these multi-dimensional

JOMH
Journal of Men's Health

TABLE 2. Eigenvalues, individual and cumulative when the target data are the mortality rate and life expectancy at birth, respectively. The raw data come from the website of the World Bank (https://data.worldbank.org) and the website of SCImago (https://www.scimagojr.com).

	Mortality rate, adult male (per 1,000 adults)			Life expectancy at birth, male (years)		
Components	Eigenvalue	Individual, %	Cumulative, %	Eigenvalue	Individual, %	Cumulative, %
1	8.606	95.626	95.626	11.036	84.890	84.890
2	0.229	2.545	98.171	1.565	12.039	96.928
3	0.116	1.289	99.459	0.229	1.763	98.691
4	0.025	0.273	99.732	0.094	0.727	99.417
5	0.011	0.127	99.859	0.049	0.376	99.793
6	0.007	0.080	99.939	0.009	0.066	99.859
7	0.003	0.032	99.971	0.008	0.063	99.922
8	0.002	0.022	99.993	0.006	0.043	99.965
9	0.001	0.007	100.000	0.002	0.015	99.980
10				0.001	0.011	99.991
11				0.001	0.006	99.997
12				0.000	0.003	99.999
13				0.000	0.001	100.000

TABLE 3. The loading when the target data are the mortality rate and the life expectancy at birth, respectively. The raw data come from the website of the World Bank (https://data.worldbank.org) and the website of SCImago (https://www.scimagojr.com).

	Mortality rate, adult (male) (per 1,000 adults)	Life expectancy at birth, (male) (years)	
Variables	The first principal component	The first principal component	The first principal component
School enrollment, tertiary (% gross)		0.805	0.579
School enrollment, tertiary, female (% gross)		0.844	0.515
Adjusted savings: education expenditure (current US$)		0.855	0.505
GDP per capita (current US$)	0.992	0.991	-0.113
Adjusted net national income per capita (current US$)		0.850	0.515
Documents in Health Professions	0.895	0.875	-0.182
Documents in Immunology and Microbiology	0.997	0.986	-0.157
Documents in Medicine	0.995	0.978	-0.186
Documents in Neuroscience	0.993	0.966	-0.235
Documents in Nursing	0.991	0.960	-0.251
Documents in Pharmacology Toxicology Pharmaceutics	0.995	0.975	-0.198
Documents in Psychology	0.963	0.900	-0.400
Documents in Veterinary	0.975	0.965	-0.146

data into one-dimensional data, then the results of calculating the target information must be independent. Therefore, there is no way to compare the comprehensive conclusion of multidimensional information collection (input information) with the information collected (input information) at each dimension. Factor analysis is a technology to extract common factors of variable groups. It is just right for solving the above problems.

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy marks the proportion of variance caused by underlying factors. The value of the Kaiser-Meyer-Olkin Measure of Sampling Adequacy was obtained as 0.876 for the mortality rate, and 0.845 for the life expectancy at birth. These values are higher than 0.5 and close to 1, which implies that the principal component analysis is significant. At the same time, all *P*-values of Bartlett's test of sphericity tests are 0.000, which reads there is a significant correlation between variables, that is, the null hypothesis that the correlation matrix is the identity matrix is negated.

Table 2 shows the eigenvalues, individual, and cumulative after those input data that are not significantly linear related to target data have been removed in the principal component analysis. According to statistics, those eigenvalues that are larger than 1 have been chosen as principal components in this paper [37–43]. Table 2 proves that there is one principal component for the target data being mortality rate, and there are two principal components when the target data are life expectancy at birth. The first principal component accounts for 95.6% of the total variance when the target data are the mortality rate and 84.9% of the total variance when the target data are the life expectancy at birth. And 12.0% of the total variance when the target data are the life expectancy at birth can be accounted for the second principal component. Two principal components can account for 96.9% of the total variance when the target data are life expectancy at birth. Thus, the difficulty in measuring the effects of knowledge

TABLE 4. Component score coefficient matrix when the target data are the mortality rate and the life expectancy at birth, respectively. The raw data come from the website of the World Bank (https://data.worldbank.org) and the website of SCImago (https://www.scimagojr.com).

	Mortality rate, adult (male) (per 1,000 adults)	Life expectancy at birth, (male) (years)	
Variables	The first principal component	The first principal component	The second principal component
School enrollment, tertiary (% gross)		-0.157	0.343
School enrollment, tertiary, female (% gross)		-0.130	0.312
Adjusted savings: education expenditure (current US$)		-0.126	0.307
GDP per capita (current US$)	0.115	0.115	-0.006
Adjusted net national income per capita (current US$)		-0.130	0.312
Documents in Health Professions	0.104	0.132	-0.048
Documents in Immunology and Microbiology	0.116	0.131	-0.029
Documents in Medicine	0.116	0.141	-0.045
Documents in Neuroscience	0.115	0.159	-0.071
Documents in Nursing	0.115	0.164	-0.079
Documents in Pharmacology Toxicology Pharmaceutics	0.116	0.146	-0.051
Documents in Psychology	0.112	0.216	-0.160
Documents in Veterinary	0.113	0.125	-0.025

[35] and the complicated relationship [56] can be analyzed by using the principal components.

The loading for the variable is shown in Table 3. The order from the largest loading to the smallest loading for the first principal component for the target data being the mortality rate is documents in neuroscience, documents in nursing, documents in veterinary, documents in pharmacology toxicology pharmaceutics, GDP per capita, documents in psychology, and documents in medicine. Table 3 also shows that the order from the largest loading to the smallest loading for the first principal component when the target data are the life expectancy at birth is GDP per capita, documents in immunology and microbiology, documents in medicine, documents in pharmacology toxicology pharmaceutics, documents in neuroscience, documents in veterinary, documents in nursing, documents in psychology, documents in health professions, adjusted savings: education expenditure, adjusted net national income per capita, school enrollment at tertiary education for female, and school enrollment, at tertiary education. The order from the largest loading to the smallest loading for the second principal component when the target data are the life expectancy at birth is school enrollment at tertiary education, adjusted net national income per capita, school enrollment at tertiary education for female, adjusted savings: education expenditure, GDP per capita, documents in veterinary, documents in immunology and microbiology, documents in health professions, documents in medicine, documents in pharmacology toxicology pharmaceutics, documents in neuroscience, documents in nursing, documents in psychology. Every loading of medical knowledge in different disciplines in the first principal components for target data being both the mortality rate and the life expectancy at birth are very high. The high loading of education expenditure for target data being the life expectancy at birth agrees well with the conclusion that both the regional growth [57, 58], and public health can be improved by education [28, 29, 53]. The loading of the number of documents in different medical disciplines is comparable with education expenditure. This is because increasing knowledge and developing knowledge can improve public health [25, 26].

Table 4 shows the component score coefficients. The order from the largest component score coefficient to the smallest loading for the first principal component for the target data being the mortality rate is documents in neuroscience, documents in nursing, documents in veterinary, GDP per capita, documents in pharmacology toxicology pharmaceutics, documents in psychology, and documents in medicine. Table 4 also demonstrates that the order from the largest component score coefficient to the smallest loading for the first principal component is documents in psychology, documents in nursing, documents in neuroscience, documents in pharmacology toxicology pharmaceutics, documents in medicine, documents in health professions, documents in immunology and microbiology, documents in veterinary, GDP per capita, adjusted savings: education expenditure, school enrollment at tertiary education for female, adjusted net national income per capita, and school enrollment at tertiary education. The order from the largest component score coefficient to the smallest loading for the second principal component is school enrollment at tertiary education, school enrollment at tertiary education for female, adjusted net national income per capita, adjusted savings: education expenditure, GDP per capita, documents in veterinary, documents in immunology and microbiology, documents in medicine, documents in health professions, documents in pharmacology toxicology pharmaceutics, documents in neuroscience, and documents in nursing, documents in psychology. Component score coefficients of medical knowledge in different disciplines in the first principal components when the target data are the mortality rate or the life expectancy at birth are high. The high component score coefficient of education expenditure supports that education can exert a very high impact on regional growth [57, 58] and an impact on health

TABLE 5. Multivariate multiple linear regression analysis results when the target data are the mortality rate and the life expectancy at birth, respectively. The raw data come from the website of the World Bank (https://data.worldbank.org) and the website of SCImago (https://www.scimagojr.com).

		Before stepwise regression analysis					
	Variable	Unstandardized Coefficients		Standardized coefficients	*t*	*Sig.*	R^2
		B	Standard Error	β			
Mortality rate, adult, male (per 1,000 male adults)			regression analysis				
	Intercept	0.000	0.116		0.000	1.000	0.693
	PC1	-0.832	0.118	-0.832	-7.045	0.000	
Life expectancy at birth, male (years)			Before stepwise regression analysis				
	Intercept	0.000	0.105		0.000	1.000	
	PC1	0.740	0.107	0.740	6.896	0.000	0.758
	PC2	0.459	0.107	0.459	4.281	0.000	
			After stepwise regression analysis				
	Intercept	0.000	0.140		0.000	1.000	0.547
	PC1	0.740	0.143	0.740	5.158	0.000	

TABLE 6. The weight oriented different targets that are the mortality rate and the life expectancy at birth, respectively. The raw data come from the website of the World Bank (https://data.worldbank.org) and the website of SCImago (https://www.scimagojr.com).

	Mortality rate, adult, male (per 1,000 adults)	Life expectancy at birth, male (years)	
Variables	The first principal component	The first principal component	The second principal component
School enrollment, tertiary (% gross)		-0.11618	0.157437
School enrollment, tertiary, female (% gross)		-0.0962	0.143208
Adjusted savings: education expenditure (current US$)		-0.09324	0.140913
GDP per capita (current US$)	-0.09568	0.0851	-0.00275
Adjusted net national income per capita (current US$)		-0.0962	0.143208
Documents in Health Professions	-0.08653	0.09768	-0.02203
Documents in Immunology and Microbiology	-0.09651	0.09694	-0.01331
Documents in Medicine	-0.09651	0.10434	-0.02066
Documents in Neuroscience	-0.09568	0.11766	-0.03259
Documents in Nursing	-0.09568	0.12136	-0.03626
Documents in Pharmacology Toxicology Pharmaceutics	-0.09651	0.10804	-0.02341
Documents in Psychology	-0.09318	0.15984	-0.07344
Documents in Veterinary	-0.09402	0.0925	-0.01148

promotion [28, 29, 53]. Most of the component score coefficients of the number of documents in different medical disciplines in China are comparable with education expenditure and adjusted net national income per capita. This might be because both education and knowledge are the most powerful engine for economic development [59], accelerated sustainable economic growth [2], increasing knowledge, and developing knowledge can improve public health [25, 26].

4. Discussion

According to the above discussion, principal components after those input data that are not significantly linear related to target data have been removed from the principal component analysis have been selected for multiple linear regressions with both the mortality rate and the life expectancy at birth.

Table 5 demonstrates the results of multivariate multiple linear regression analysis results based on principal component score. 69.3% of the variation in the normalized male mortality rate (adult) of China could be explained by both the first principal components. 54.7% of the variation in the normalized male life expectancy at birth in China could be explained by the first principal component, which is determined from the stepwise regression analysis. 75.8% of the variation in the normalized male life expectancy at birth in China could be explained by both the first and second principal components. The above results reveal that the first principal component gives the most contribution to both the mortality rate and the life expectancy at birth in China. Since both *P* values are 0.000 is less than 0.05, it is concluded that the regression model (with several predictors included) significantly predicts male mortality rate and male life expectancy at birth in China. On the other hand, the *P*-values for the principal components shown in Table 5 are less than 0.05, which means that principal components account for a significant amount of unique variance in male mortality rate and male life expectancy at birth of China. The *R*-values can show that the collected parameters better explain male life expectancy than that for male mortality rate and more data need to be considered in predicting male mortality rate

and male life expectancy at birth in China.

These results of multiple linear regressions further support the well-known conclusion that education can exert a very high impact on both regional growths [2, 57–59], and health promotion [25, 26, 28, 29, 53], lack of knowledge can cause serious health problems [27]. Knowledge is necessary for health benefits [22]. It is because of the function of knowledge: making proper decisions and realizing actions [2], and health can be improved by making better health decisions [29]. Note that the above conclusions are mainly about the effect of knowledge on health. There is no specific discussion about the role of knowledge, especially medical knowledge in different fields on men's health. Different from these research results, this paper is to study the impact of published papers in various disciplines of Medicine on men's health. To our knowledge, this is the first study to discuss the impact of the medical publication in various disciplines on male mortality rate and male life expectancy at birth in China by using the regression method based on the principal component analysis. The language of medical papers published by China not only focuses on Chinese, thus which may not be a good reflection of the whole Chinese population because many people do not read English papers.

To compare the weight of input variables oriented a target, the weight of the i-th input variable oriented a target can be defined as

$$W_i^j = S_i^j \times B^j \tag{3}$$

where B^j is the regression coefficient between the input variables and the j-th target data, and S_i^j is the component score coefficient of the i-th input variable oriented a target.

Table 6 shows the results of evaluating the role of the articles in medical disciplines oriented different targets by using Eqn. 3. The order for the maximum contribution to the minimum contribution to decreasing the male mortality rate (adult) in the first principal component is documents in medicine, GDP per capita, documents in pharmacology toxicology pharmaceutics, documents in psychology, documents in neuroscience, documents in nursing, and documents in veterinary. The order for the maximum contribution to the minimum contribution to increasing male life expectancy at birth in the first principal component documents in psychology, documents in nursing, documents in neuroscience, documents in pharmacology toxicology pharmaceutics, documents in medicine, documents in health professions, documents in immunology and microbiology, documents in veterinary, GDP per capita, adjusted savings: education expenditure, school enrollment at tertiary education for female, adjusted net national income per capita, and school enrollment at tertiary education.

5. Conclusions

After those input data that are not significantly linear related to target data have been removed in the principal component analysis, principal components have been selected for multiple linear regressions with both the mortality rate and the life expectancy at birth. 69.3% of the variation in the normalized male mortality rate (adult) of China could be explained by both the first principal components. 54.7% of the variation in the normalized male life expectancy at birth in China could be explained by the first principal component, which is determined from the stepwise regression analysis. 75.8% of the variation in the normalized male life expectancy at birth in China could be explained by both the first and second principal components. The above results reveal that the first principal component gives the most contribution to both the mortality rate and the life expectancy at birth in China. These results of multiple linear regressions further support the well-known conclusion reported in the literature. These findings further support the well-known conclusion that education can exert a very high impact on both regional growth and health promotion, and lack of knowledge can cause serious health problems. The proposed method can be treated as a new method introduced to study the correlation between scholarly publishing and health and well-being. In conclusion, principal component analysis is valid in the study of the correlation between scholarly publishing and health and well-being.

Author contributions

R. conceived and designed the experiments; R. performed the experiments; R. analyzed the data; R. wrote the paper.

Ethics approval and consent to participate

Not applicable.

Acknowledgment

Thanks to all the peer reviewers and editors for their opinions and suggestion.

Funding

Not applicable.

Conflict of interest

The authors declare no competing interests.

References

[1] Kinzig AP, Carpenter S, Dove M, Michael M, Heal G, Levi, S, *et al.* Nature and society: An imperative for integrated environmental research. In Executive summary of a workshop sponsored by NSF, Developing a Research Agenda for Linking Biogeophysical and Socioeconomic Systems (p. 72). Tempe, Arizona. 2000.

[2] Sarbu M, Doinea O, Mangra M. Knowledge based economy-the basis for insuring a sustainable development. Annals of the University of Petrosani, Economics. 2009; 9: 6.

[3] Jacobson N, Butterill D, Goering P. Development of a framework for knowledge translation: understanding user context. Journal of Health Services Research & Policy. 2003; 8: 94-99.

[4] Wensing M, Grol R. Knowledge translation in health: how implementation science could contribute more. BMC Medicine. 2019; 17: 88.

[5] Smylie J, Kaplan-Myrth N, McShane K. Indigenous knowledge translation: baseline findings in a qualitative study of the pathways of health knowledge in three indigenous communities in Canada. Health Promotion Practice. 2009; 10: 436-446.
[6] Tetroe JM, Graham ID, Foy R, Robinson N, Eccles MP, Wensing M, *et al.* Health research funding agencies' support and promotion of knowledge translation: an international study. The Milbank Quarterly. 2008; 86: 125-155.
[7] LaRocca R, Yost J, Dobbins M, Ciliska D, Butt M. The effectiveness of knowledge translation strategies used in public health: a systematic review. BMC Public Health. 2012; 12: 751.
[8] Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, *et al.* Lost in knowledge translation: time for a map? The Journal of Continuing Education in the Health Professions. 2006; 26: 13-24.
[9] Estabrooks CA, Thompson DS, Lovely JJE, Hofmeyer A. A guide to knowledge translation theory. The Journal of Continuing Education in the Health Professions. 2006; 26: 25-36.
[10] McKibbon KA, Lokker C, Wilczynski NL, Ciliska D, Dobbins M, Davis DA, *et al.* A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? Implementation Science. 2010; 5: 16.
[11] Koens L, Harkema B, Faasse P. Making knowledge work: the function of public knowledge organizations in the Netherlands. Research outside the Academy. 2019; 44: 71-87.
[12] Barwick MA, Peters J, Boydell K. Getting to uptake: do communities of practice support the implementation of evidence-based practice? Journal of the Canadian Academy of Child and Adolescent Psychiatry. 2009; 18: 16-29.
[13] Noia JD, Schwinn TM, Dastur ZA, Schinke SP. The relative efficacy of pamphlets, CD-ROM, and the Internet for disseminating adolescent drug abuse prevention programs: an exploratory study. Preventive Medicine. 2003; 37: 646-653.
[14] Dobbins M, Hanna SE, Ciliska D, Manske S, Cameron R, Mercer SL, *et al.* A randomized controlled trial evaluating the impact of knowledge translation and exchange strategies. Implementation Science. 2009; 4: 61.
[15] Forsetlund L, Bradley P, Forsen L, Nordheim L, Jamtvedt G, Bjørndal A. Randomised controlled trial of a theoretically grounded tailored intervention to diffuse evidence-based public health practice [ISRCTN23257060]. BMC Medical Education. 2003; 3: 2.
[16] Hutchinson JR, Huberman M. Knowledge dissemination and use in science and mathematics education: a literature review. Journal of Science Education and Technology. 1994; 3: 27-47.
[17] Huberman M. Linkage between researchers and practitioners: a qualitative study. American Educational Research Journal. 1990; 27: 363-391.
[18] Rogers EM. Diffusion of Innovations. 4th edition. Free Press. 2010.
[19] Greenhalgh T, Robert G, Bate P, Macfarlane F, Kyriakidou O. Diffusion of innovations in health service organisations: a systematic literature review. Wiley. 2007; 1-316.
[20] Albrecht L, Archibald M, Arseneau D, Scott SD. Development of a checklist to assess the quality of reporting of knowledge translation interventions using the Workgroup for Intervention Development and Evaluation Research (WIDER) recommendations. Implementation Science. 2013; 8: 52.
[21] Tapfuma M, Hoskins R. Visibility and accessibility of indigenous knowledge on open access institutional repositories at universities in Africa. In Information Resources Management Association (eds) Digital libraries and institutional repositories (pp. 454-472). Hershey PA: IGI Global. 2020.
[22] Cale L, Harris J, Hooper O. Debating health knowledge and health pedagogies in physical education. In Capel S., Blair R. (eds) Debates in physical education (pp. 256-277). New York: Routledge. 2019.
[23] Durkalec A, Furgal C, Skinner MW, Sheldon T. Climate change influences on environment as a determinant of Indigenous health: Relationships to place, sea ice, and health in an Inuit community. Social Science & Medicine. 2015; 136-137: 17-26.
[24] Taş F, Kocaöz S, Çirpan R. The effect of knowledge and health beliefs about colorectal cancer on screening behaviour. Journal of Clinical Nursing. 2019; 28: 4471-4477.
[25] Shawe J, Patel D, Joy M, Howden B, Barrett G, Stephenson J. Preparation for fatherhood: a survey of men's preconception health knowledge and behaviour in England. PLoS ONE. 2019; 14: e0213897.
[26] Silva N, Mekaro KS, Santos RIO, Uehara S. Knowledge and health promotion practice of Family Health Strategy nurses. Revista Brasileira de Enfermagem. 2020; 73: e20190362.
[27] Hansen RP, Vedsted P, Sokolowski I, Søndergaard J, Olesen F. Time intervals from first symptom to treatment of cancer: a cohort study of 2,212 newly diagnosed cancer patients. BMC Health Services Research. 2011; 11: 284.
[28] Grossman M. Chapter 10 education and nonmarket outcomes. In Hanushek E., Welch F. (eds) Handbook of the economics of education (pp. 577-633). Elsevier. 2006.
[29] Kemptner D, Jürges H, Reinhold S. Changes in compulsory schooling and the causal effect of education on health: evidence from Germany. Journal of Health Economics. 2011; 30: 340-354.
[30] Ghodsbin F, Zare M, Jahanbin I, Ariafar A, Keshavarzi S. A survey of the knowledge and beliefs of retired men about prostate cancer screening based on health belief model. International Journal of Community Based Nursing and Midwifery. 2014; 2: 279-285.
[31] Alduraywish SA, Altamimi LA, Aldhuwayhi RA, AlZamil LR, Alzeghayer LY, Alsaleh FS, *et al.* Sources of health information and their impacts on medical knowledge perception among the Saudi Arabian population: cross-sectional study. Journal of Medical Internet Research. 2020; 22: e14414.
[32] De Jesus M. Mutuality at the center: health promotion with Cape Verdean immigrant women. Ethnicity & Health. 2009; 14: 45-59.
[33] Massey PM, Prelip M, Calimlim BM, Quiter ES, Glik DC. Contextualizing an expanded definition of health literacy among adolescents in the health care setting. Health Education Research. 2012; 27: 961-974.
[34] Fadaiyan Arani E, Amin Shokravi F, Tavakoli Ghouchani H. The relationship between health literacy and knowledge in rural patients with type 2 diabetes mellitus in 2016. Health Education and Health Promotion. 2017; 5: 19-32.
[35] Fagerberg J, Srholec M. Knowledge, capabilities, and the poverty trap: the complex interplay between technological, social, and geographical factors. In Meusburger P., Glückler J., el Meskioui M. (eds) Knowledge and the economy (pp. 113-137). Springer. 2013.
[36] Cho YI, Lee SD, Arozullah AM, Crittenden KS. Effects of health literacy on health status and health service utilization amongst the elderly. Social Science & Medicine. 2008; 66: 1809-1816.
[37] Fox WP. Mathematical modeling for business analytics. Routledge & CRC Press. 2017.
[38] Bowden R. The information theory of comparisons. Springer. 2018.
[39] Mukherjee SP, Sinha BK, Chattopadhyay AK. Statistical methods in social science research. Springer. 2018.
[40] Melnik R. Mathematical and computational modeling: with applications in natural and social sciences, engineering, and the arts. Wiley. 2015.
[41] Grimm LG. Statistical applications for the behavioral sciences. Wiley. 1993.
[42] Strawinska-Zanko U, Liebovitch LS. Mathematical modeling of social relationships. Springer. 2018.
[43] Warne RT. Statistics for the social sciences: a general linear model approach. Cambridge University Press. 2017.
[44] Wang L, Fu X. Data mining with computational intelligence. Springer. 2006.
[45] Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences. 2016; 374: 20150202.
[46] Zhao Z, Fan X, Qi Y, Zhai Y. Multi-angle insulator recognition method in infrared image based on parallel deep convolutional neural networks. In Yang J. *et al.* (eds) Computer Vision. CCCV 2017. Communications in computer and information science (pp. 303-314). Springer. 2017.
[47] Ali A, Shang J, Saif U. Socio-economic impact of CPEC on agricultural

productivity of Pakistan: a principal component analysis. International Journal of Food and Agricultural Economics. 2018; 6.

[48] Begdache L, Kianmehr H, Sabounchi N, Marszalek A, Dolma N. Principal component regression of academic performance, substance use and sleep quality in relation to risk of anxiety and depression in young adults. Trends in Neuroscience and Education. 2019; 15: 29-37.

[49] Fadhel S, Delpha C, Diallo D, Bahri I, Migan A, Trabelsi M, *et al.* PV shading fault detection and classification based on I-V curve using principal component analysis: application to isolated PV system. Solar Energy. 2019; 179: 1-10.

[50] Mehmet M. Multivariate multiple regression analysis based on principal component scores to study relationships between some pre- and post-slaughter traits of broilers. TarıM Bilimleri Dergisi. 2011; 17: 77-83. (In Turkish)

[51] Papi M, Caracciolo G. Principal component analysis of personalized biomolecular corona data for early disease detection. Nano Today. 2018; 21: 14-17.

[52] Sousa S, Martins F, Alvimferraz M, Pereira M. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environmental Modelling & Software. 2007; 22: 97-103.

[53] Rosenzweig MR, Schultz TP. Schooling, information and nonmarket productivity: contraceptive use and its effectiveness. International Economic Review. 1989; 30: 457-477.

[54] Brenner MH. Commentary: economic growth is the basis of mortality rate decline in the 20th century-experience of the United States 1901-2000. International Journal of Epidemiology. 2005; 34: 1214-1221.

[55] Pritchett L, Summers LH. Wealthier is healthier. The Journal of Human Resources. 1996; 31: 841.

[56] Mokyr J. A Culture of Growth. The Origins of the Modern Economy. Princeton University Press. 2017.

[57] Romão J, Neuts B. Territorial capital, smart tourism specialization and sustainable regional development: experiences from Europe. Habitat International. 2017; 68: 64-74.

[58] Sterlacchini A. R&D, higher education and regional growth: uneven linkages among European regions. Research Policy. 2008; 37: 1096-1107.

[59] Glückler J, Meusburger P, Meskioui ME. Introduction: knowledge and the geography of the economy. In Meusburger P., Glückler J., el Meskioui M. (eds) Knowledge and the economy (pp. 3-14). Springer. 2013.