Review Article

Biomathematical model study on the opioid crisis in America

Abstract

In the process of this paper, we started with the data from 2010 to 2017 provided by NFLIS concerned with opioid crisis in the US. After sorting and analyzing the panel data, we decided to transform the derived data and then model the panel data, cross-section data and time series data respectively. Next, we consider how to objectively select a large number of socio-economic data and indicators, and then establish a model that can reflect two different databases at the same time, so that the model can be combined with some indicators. Finally, we propose a feasible strategy to resist the opioid crisis.

Keywords: dynamic panel data model, systematic cluster analysis, correlation analysis, weighted principal component analysis

Bin Zhao,[1] Xia Jiang,[2] Jinming Cao,[3] Kuiyun Huang,[1] Jingfeng Tang[4]
[1]School of Science, Hubei University of Technology, China
[2]Hospital, Hubei University of Technology, China
[3]School of Information and Mathematics, Yangtze University, China
[4]National "111" Center for Cellular Regulation and Molecular Pharmaceutics, Hubei University of Technology, China

Correspondence: Dr. Bin Zhao, School of Science, Hubei University of Technology, Wuhan, Hubei, China, Tel/Fax +86 130 2851 7572, Email zhaobin835@nwsuaf.edu.cn

Introduction

The United States opioid epidemic is a nationwide public health crisis. Opioids are prescription drugs, and three-thirds of the deaths from opioids die each year from opioid prescription drugs. Heroin is one of the most highly dependent substances in opioids. It is more addictive than any other drug and is loved by addicts. Although heroin mortality is high, it is still difficult to control.[1] Since President Nixon launched a large-scale anti-drug operation in the late 1960s, successive US administrations have made unremitting efforts in combating drug smuggling and controlling the spread of drugs. They have also set up the Narcotics Bureau to kill drugs from the root causes. According to statistics, an average of 40 people die every day in the United States due to overdose of opioids. This number has tripled since 1999.[2] The National Drug Use and Health Survey (NSDUH) of the US Drug Abuse and Mental Health Service Management (SAMHSA) shows that in 2016, more than 11 million people in the United States abused opioid prescription drugs and nearly 1 million people used heroin. All of the above indications indicate that the use of opioids in the United States has caused serious social problems. The real need for the United States to face up is the growing gap between the rich and the poor, the inability to make ends meet, and the low-education employment opportunities.

Methods

We discuss the extent of opioids flooding, and establishes a dynamic panel data model for characterization and prediction of crisis, using generalized moment estimation and time series analysis to solve the model; the grey relational analysis is carried out to judge whether the use of opioids is related to population data, and the principal component evaluation model is established to verify the results; the linear programming model was analyzed using the sensitivity analysis for the validity of the test strategy. First, a multi-dimensional descriptive statistical analysis of the data, and found the geographical distribution of opioids. The data was organized into panel data. On the one hand, the dynamic panel data model was established and the

parameters were estimated by generalized moments. It was concluded that the heroin should first appeared in 1910, OH-HAMILTON and synthetic opioids first appeared in 1939 PA-PHILADELPHIA. On the other hand, the Hierachical Cluster based on the panel data of "absolute quantity", "fluctuation", "skewness", "kurtosis" and "trend" feature extraction is used to find out the five counties that need the most concern in the US. And time Series Analysis was used to find the year when these counties reached the drug threshold and the threshold level was obtained by the dynamic panel data model. For example, the threshold for the number of synthetic opioid cases in OH-CUYAHOGA in 2018 was 6783.

To judge whether the use of opioids is related to US population data, use gray correlation analysis to find the ratio of the number of heroin and synthetic opioid cases in the county to the total number of identified substances. The correlation between the two and the selected first-level indicators from the population data is greater than 0.5, indicating that the degree of association between them is greater. To verify the five counties that may cause the US to panic, the entropy weight method is used to select 16 indicators, and the indicators in the NFLIS are combined to establish a comprehensive evaluation system for the degree of opioid flooding and principal component evaluation model. The comprehensive weighted scores were used as the class of opioid flooding scores, and the 461 counties in 8 years were distributed according to the frequency distribution of F values, and the degree of opioid abuse was divided into three levels: severe, general and lower. We propose the strategies for combating the opioid crisis: the US can stipulate that all people must have completed the 12th grade compulsory education when they are 25 years old. To test the effectiveness of the strategy and determine the range of important parameters, a sensitivity analysis of linear programming was used. Taking min F1 as the objective function, a constraint condition is formed between the 20 indicators, and the parameter range (c, k) of each index is obtained by local sensitivity analysis. The obtained parameter range is brought into the first principal component expression, and it is determined whether the parameter range is valid according to the level of the F1. In the end, the parameters of high

school education, university but no degree and university degree and above are correct (0,0.2637), (0.2615,1), (0.2615,1), and the flood levels of the four counties are correct. Has been reduced to a lower level, only Hamilton County, Ohio's hazard level reduced to a general conclusion, which shows that the strategy is effective.

Biomathematical model establishment and solution

Biomathematical modelling

Multidimensional descriptive statistical analysis model

This study investigates the data concerned with opioid crisis from 461 counties in the five states from 2010 to 2017, with a total of 24,063 samples and 61 substance name.

Counties: The data can be sorted out. In the 461 counties, not every county has an incident report every year. The reason may be that the data of the current year is difficult to obtain, or the data of the year is 0, so it is omitted. However, we believe that either case can indicate that the county's drug abuse has not reached a serious level. Combined with the data of each county, the counties with missing data have fewer drugs in the few years with data, close to zero, so we fill the value of the drug that was missing in the county from zero.

Substance Name: The name of the substance identified in the analysis contains 47 synthetic opioids and 13 non-synthetic opioids, these 13 non-synthetic drugs are Codeine, Dihydrocodeine, Acetylcodeine, Acetyldihydrocodeine, Morphine, Heroin, Hydromorphone, Oxycodone, Oxymorphone, Buprenorphine, Hydrocodone, Nalbuphine, Dihydromorphone.

YYYY&FIPS_Combined: The time and county code data are all integers, and the distance between the data is the same, so the two columns of data are logarithmically transformed to make the data better visualized, and the difference between the two columns is small, especially year. Therefore the logarithmic transformation is performed with a base of 1.1.

The overall trend of the number of drugs in the five states from 2010 to 2017 and the overall trend of the number of heroin were analyzed. The resulting bar chart is shown in Figures 1 & 2. It can be seen intuitively from Figure 1 that the total number of drugs in KY and OH states is far greater than the other three states. Among them, the OH state has increased year by year, the PA state has decreased year by year, the VA state has fluctuated greatly, and the KY state and the WV state have stabilized at a lower value. As can be seen from Figure 2, the number of heroin in these five regions increased first and then decreased over time. The turning point is probably in 2015, and the number of heroin in OH and PA states is much higher than in the other three states. Therefore, PA State and OH State are the targets of key observations. Further analysis of the change in the proportion of the substance identified in the analysis from 2010 to 2017, the resulting percentage of the accumulated area is shown in Figure 3. The greater the proportion of color in the percentage stacked graph, the greater the proportion of the substance in the analysis. It can be seen from Figure 3 that heroin (dark brown part of Figure 3) accounts for the most, nearly half. Followed by Oxycodone (light grey), the proportion of other substances is much smaller than heroin. Continue to observe the geographical distribution of the number of heroin, as shown in Figure 4. It can be seen from Figure 4 that heroin is concentrated in five counties, and the codes according to the distribution order (FIPS) are 39035, 39061, 39113, 42003, 42101.
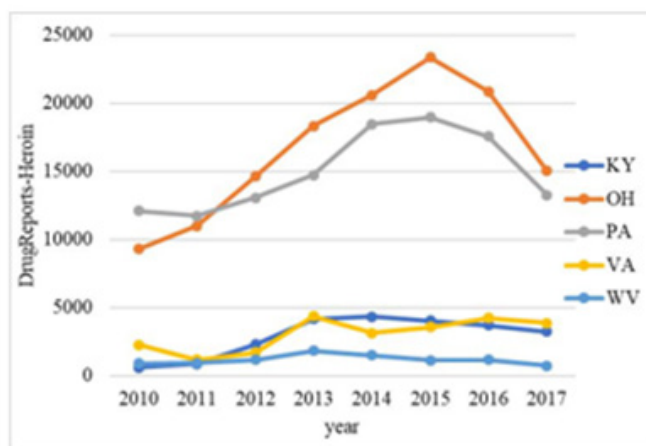

Figure 1 Changes in the number of drugs in five states.
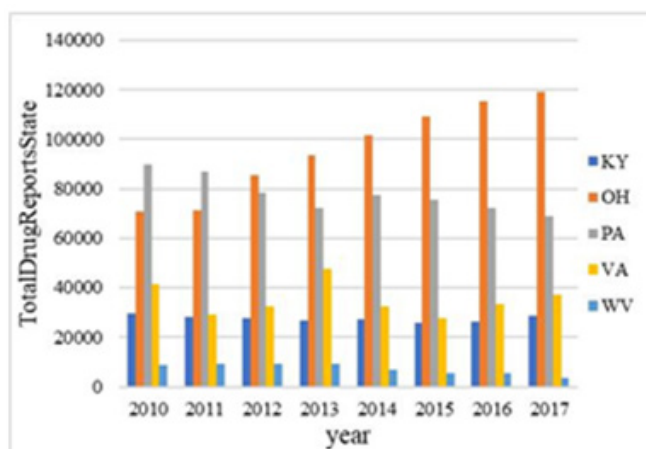

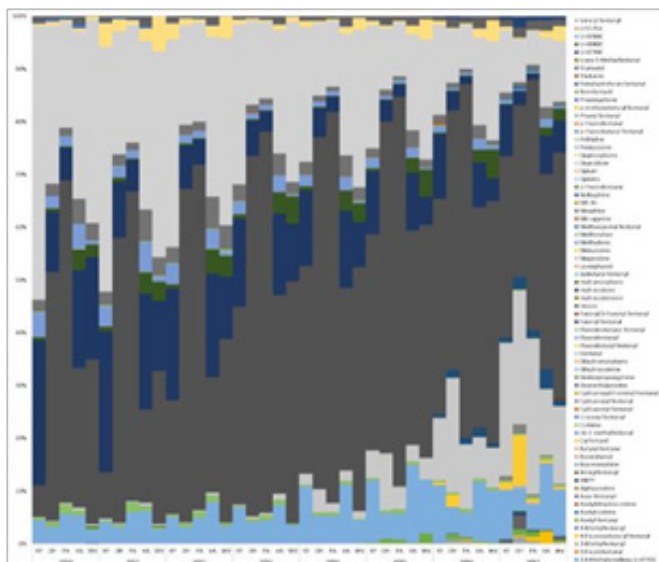**Figure 2** Changes in the number of heroin.

Figure 3 percentage stacked column chart.

The establishment of dynamic panel data model

The data given in the title is multi-indicator panel data. In order to facilitate the observation of indicators, the data is organized into

the form of Table 1. Strictly speaking, it should be represented by a three-dimensional table. For ease of understanding and explanation, the following two tables are still used. As shown in Table 1, there are a total of N samples in the study. Each sample has T records and p indicators per period. Then the value of the j-th indicator of the sample i in the t-th period is $X_{ij}(t)$, where $i=1,2\cdots N j=1,2\cdots p t=1,2\cdots T$, the difference between this table and the simple two-dimensional table is that it contains three-dimensional information such as time, sample and indicator.

Data preprocessing

The data of time and county code are integers, and the distance between the data is the same, so the two columns of data are logarithmically transformed to make the data better visual, and the difference between the two columns is small, especially time. Therefore, the logarithmic transformation is performed with a base of 1.1.

Figure 4 Geographical distribution of opioids.

Table 1 Multiple indicator panel data

Time	1			...	t			...	T		
Sample	$X_1...X_j...X_p$			...	$X_1...X_j...X_p$			...	$X_1...X_j...X_p$		
1	$X_{11}(1)$	...	$X_{1p}(1)$	...	$X_{11}(t)$	...	$X_{1p}(t)$	...	$X_{11}(T)$	...	$X_{1p}(T)$
2	$X_{21}(1)$	...	$X_{2p}(1)$	...	$X_{21}(t)$	...	$X_{2p}(t)$	...	$X_{21}(T)$	...	$X_{2p}(T)$
...	...	...	...	...	...	...	...	...	...	...	...
i	$X_{i1}(1)$	...	$X_{ip}(1)$	...	$X_{i1}(t)$	...	$X_{ip}(t)$	...	$X_{i1}(T)$	...	$X_{ip}(T)$
...	...	...	...	...	...	...	...	...	...	...	...
N	$X_{N1}(1)$	...	$X_{Np}(1)$	...	$X_{N1}(t)$	...	$X_{Np}(t)$	...	$X_{N1}(T)$	...	$X_{Np}(T)$

Dynamic panel data model establishment

Since the title requires determining the earliest position used by the specific opioid, the quantization model is prioritized to incorporate the position as a variable into the model to solve the earliest position. Therefore, the regression model is used. The initial model is as follows

$$Y=\alpha+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\mu \qquad (3.1)$$

Among them, Y represents the number of drugs identified in each county, X1 represents the code of each county, X2 represents time (yearly), and X3 represents the ratio of the number of identified drugs in each county to the total number of confirmed drug cases in the county for we believe that the ratio can reflect the degree of development of drugs to some extent. Considering the indelibility and infectivity of drugs, the drugs that lag behind the first phase should have an impact on the previous period. Therefore, the lag phase is included as an explanatory variable in the model, and the first-order lag variable is considered. Since the explanatory variables contain both time variables and regional variables, and cover both time series data and cross-section data, the dynamic panel data model is finally adopted. Since the number and proportion of drugs have zero values, the initial model (3.1) is adjusted as follows.

$$Y=\alpha+\beta_0 Y_{i(t-1)}+\beta_1\log_1 X_1+\beta_2\log_1 X_2+\beta_3 X_3+\mu \qquad (3.2)$$

where y_{it}=the number of drugs identified in the t-th year of the i-th county.

Model solving and analysis

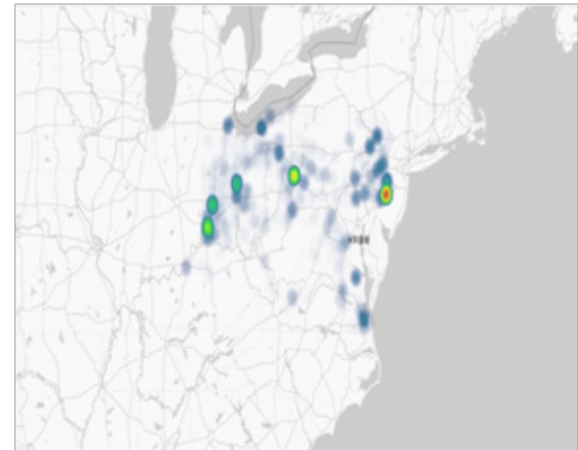In the dynamic panel data model, due to the existence of the lag-interpreted variable, it is possible that the explanatory variable is related to the random error term, so that the estimators obtained by using OLS and GLS are biased and non-uniform. Ahn and Schmidt (1995) and Judson and Oewn (1999) used generalized moments (GMM) to study the parameter estimation of the dynamic panel data model, the statistical properties of the estimates and the model checking methods.[3] The core idea of GMM estimation is to use tool variables to generate corresponding moment conditions.

According to the estimation idea of GMM, the model (3.2) is estimated by EVIEWS software, and the hysteresis order is determined according to whether the t-test of the parameter estimation has robustness. The result is as follows.

Section 1: Heroin

$$Y_{it}=84481.93+0.9934Y_{i(t-1)}-0.4126\log_{1.1}X_1-105.856\log_{1.1}X_2+75.018X_3+\mu$$

$$R^2=0.849(216.4359)^{***}(-1.0367)(-7.2537)^{***}(5.7406)^{***}$$

It can be seen from the above formula that the adjusted R square is 0.849, and the fitting effect is general, and X_1 does not pass the 10% t test. However, the correlation matrix shows that there is no multicollinearity between X_1 and other explanatory variables, so the data is adjusted to raw data. The model is improved as follows.

$$Y_{it}=548.502+0.9935Y_{i(t-1)}-\frac{21323825.669}{X_1}-75.339X_2-0.00137X_3^2+\mu$$

$$AQF(F_{ij})=\frac{\sum\limits_{t=1}^{T}X_{ij}^{*}(t)}{T}$$

Among them, the t-statistic is in the brackets, *** indicates that it is significant at the level of 0.05, and * indicates that it is significant at the level of 0.1. Except for the dynamic panel data model, the county code is significant at the level of 0.1, and other estimates are significant at the 0.05 level, indicating that the regression effect is good.

Section 2: Synthetic opioid

$$Y_{it} = -16277.25 + 2.0342Y_{i(t-1)} - 0.126\log_t X_1 - 204.764\log_t X_2 + 39.4997X_3^2 + \mu$$

$$R^2 = 0.849(216.4359)^{***}(-1.0367)(-7.2537)^{***}(5.7406)^{***}$$

It can be seen from the above formula that the adjusted R square is 0.949, and the fitting effect is good, and each explanatory variable has passed the significant level of 0.01.

The earliest appearance of opioids must have never appeared before, and began to grow after emergence. Therefore, the panel data model can be used to make $Y_{it} = 0$, then $Y_{i(t-1)}$ and ratio are 0. To find the earliest position, we hope that the year is as small as possible. We have obtained a quantitative relationship between the area and the number of drugs, so that the year can be reduced in turn, and the corresponding county codes are obtained separately. In the case of Heroin, when the year is reduced to 1910, the corresponding county code is 39061. When the year is reduced to 1909, the county code has been reduced to four digits, which is inconsistent with the data. Therefore, 1910 is considered to be Heroin. Because of the earliest year, the earliest position corresponding to the occurrence is 39061 (OH-HAMILTON). The earliest appearance time of synthetic opioids is 1939, and the earliest position is 42101 (PA-PHILADELPHIA).

Time series model based on system clustering

Five feature extraction of panel data

Several statistics for the multidimensional indicator panel are given below, and the statistic feature extraction will use these statistics.

The mean and standard deviation of the j th indicator T period of sample i are:

$$\mu_{ij} = \frac{\sum\limits_{t=1}^{T} X_{ij}(t)}{N}, \sigma_{ij} = \left[\frac{\sum\limits_{t=1}^{T}(X_{ij}(t) - \overline{X}_{ij})^2}{N}\right]^{\frac{1}{2}}$$

Standardization of panel data

Due to the difference in the dimension and magnitude of the indicator, it will have an impact on the final analysis results. Therefore, the standardization process of he mean of $X_{ij}(t)$ is first performed, and the standardized data is set to $X_{ij}^*(t)$, and the standardized formula is Due to the difference in the dimension and magnitude of the indicator, it will have an impact on the final analysis results. Therefore, the standardization process for the mean value is first set, and the standardized data is

$$X_{ij}^*(t) = \frac{X_{ij}(t)}{\overline{X}_j} \tag{3.3}$$

Among them $\overline{X}_j = \dfrac{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} X_{ij}(t)}{NT}$ after standardization, the mean value of each indicator is 1, and the variance is

$$Var(X_j^*) = \frac{1}{NT-1}\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T}\left[\frac{X_{ij}(t)}{\overline{X}_j} - 1\right]^2 = \frac{Var(X_j)}{\overline{X}_j^2} = \left(\frac{\sigma_j}{\overline{X}_j}\right) \tag{3.4}$$

The variance of each index after such standardization is the square of the coefficient of variation of each index, which not only eliminates the influence of dimension and magnitude, but also retains the variation information of the original indicator.

Feature quantity extraction of panel data indicators

According to the extraction of the feature quantity of panel data in the literature,[4] this paper defines the feature quantity of each index during the inspection period from the aspects of development level, trend, fluctuation degree and distribution of the indicator period. For the panel dataset $\{X_{ij}^*(t)\}$, there are N samples, each sample records T, and there are p indicators in each period.

Definition 1: The jth indicator of the sample i is the full-time Absolute Quantity Feature, abbreviated as $AQF(F_{ij})$.

$$AQF(F_{ij}) = \frac{\sum\limits_{t=1}^{T} X_{ij}^*(t)}{T} \tag{3.5}$$

$AQF(F_{ij})$ is actually the mean of the jth indicator of sample i over the total period T, which reflects the absolute level of development of the jth indicator of sample i in the analysis time domain (over the entire period).

Definition 2: The jth indicator of the sample i is the full-time "Variance Feature", abbreviated as $VF(F_{ij})$, then

$$VF(F_{ij}) = \left[\frac{\sum\limits_{t=1}^{T}\left(X_{ij}^*(t) - \overline{X}_{ij}^*\right)}{T-1}\right]^{\frac{1}{2}} \tag{3.6}$$

Among them, $\overline{X}_{ij}^* = \dfrac{\sum\limits_{t=1}^{T} X_{ij}^*(t)}{T}$ is $AQF(F_{ij}), VF(F_{ij})$ in definition 1, which reflects the degree of fluctuation of the jth index of sample i over time.

Definition 3: The jth indicator of the sample i is the full-time Skewness Coefficient Feature, abbreviated as

$$SCF(F_{ij}) = \frac{\sum\limits_{t=1}^{T}\left(X_{ij}^*(t) - \overline{X}_{ij}^*\right)^3}{T\left(\sigma_{ij}^*\right)^3} \tag{3.7}$$

Where $\sigma_{ij}^* = \left[\dfrac{\sum\limits_{t=1}^{T}\left(X_{ij}^*(t) - \overline{X}_{ij}^*\right)}{T-1}\right]^{\frac{1}{2}}$ represents the standard deviation of the jth index of the sample i over the entire period,

$SCF(F_{ij})$ reflects the degree of symmetry of the jth index of the sample i over the entire period, $SCF(F_{ij}) < 0$, indicating that most of the index is located to the right of the average, $SCF(F_{ij}) < 0$, indicating that most of indicators are located to the left of the average.

Definition 4: The jth indicator of the sample i is the Kurtosis Coefficient Feature, abbreviated as $KCF(F_{ij})$.

$$KCF(F_{ij}) = \frac{\sum_{t=1}^{T}\left(X_{ij}^{*}(t) - \overline{X}_{ij}^{*}\right)}{T\left(\sigma_{ij}^{*}\right)} - 3 \qquad (3.8)$$

$KCF(F_{ij})$ reflects the sharpness of the distribution curve of the jth indicator of sample i over the entire period.

$KCF(F_{ij}) > 0$ indicates that the distribution of the index value is more dispersed than the normal distribution, and $KCF(F_{ij}) < 0$ indicates that the distribution of the index value is more concentrated

around the average value than the normal distribution.

Definition 5: The jth indicator full-time "Trend Feature" of sample i, abbreviated as $TF(F_{ij})$, the long-term trend of the $TF(F_{ij})$ indicator. If the $TF(F_{ij})$ value of the indicator is closer, it means that both indicators show the same slope change and the closer the two indicators are.

Indicator selection

According to the previous analysis of the data and the variables of the demand, feature extraction of the following indicators: Heroin, non-synthetic opioids, Synthetic opioid, opioids, Total Drug Reports County.

Extraction results

Take the Absolute Quantity Feature as an example. The final data obtained is shown in Table 2 below.

In order to visually see the data characteristics of different indicators in the time dimension between the county and the county, take the Absolute Quantity Feature as an example and make an observation chart of five indicators. The obtained line chart is shown in Figure 5.

Table 2 Absolute Quantity Feature

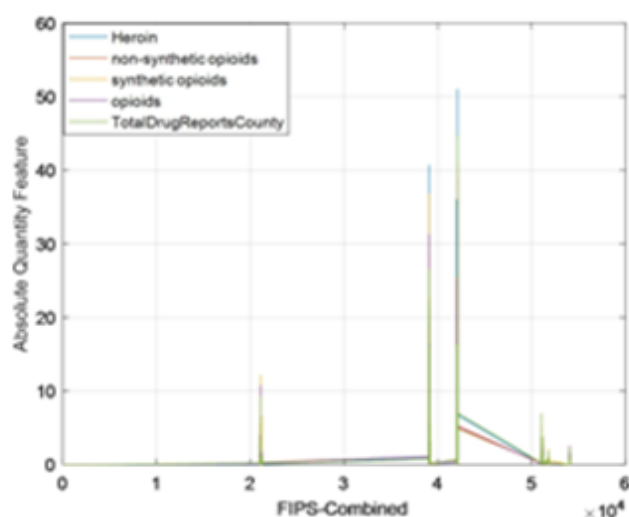|	Descriptive Statistics FIPS_ Heroin non-synthetic Combined	Heroin	non-synthetic opioids	AQF synthetic opioids	opioids	Total Drug Reports County
1	21001	0.01	0.14	0.1	0.13	0.21
2	21003	0	0.21	0.26	0.2	0.23
3	21005	0.13	0.16	0.27	0.16	0.14
…	…	…	…	…	…	…
461	54109	0	0.05	0	0.04	0.03

Figure 5 AQF.

The abscissa in Figure 5 indicates (total coding) FIPS_Combined, the ordinate indicates the AQF value corresponding to each county, and Figure 6 is a partial enlarged view of. 5. As can be seen from Figure 5 and Figure 6, the fluctuations of these indicators are similar.

Explain heroin, non-synthetic drugs, synthetic drugs, opioids, total drug counts in Total Drug Reports County. These indicators have similar development levels throughout the period from 2010 to 2017, and each county has its own characteristics (Appendix 1).
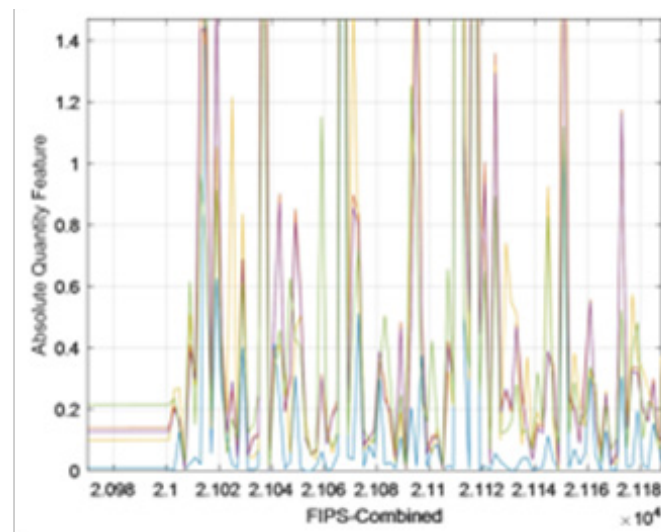

Figure 6 AQF(part).

Citation: Zhao B, Jiang X, Cao J, et al. Biomathematical model study on the opioid crisis in America. *MOJ App Bio Biomech.* 2019;3(2):34–46.
DOI: 10.15406/mojabb.2019.03.00097

Cluster data clustering results

The five characteristics extracted from the panel data index were systematically clustered with heroin and synthetic opioids. The obtained pedigree map is shown in Appendix 2. The systematic clustering results of synthetic opioids and the systematic clustering results of heroin. Consistent, see Appendix 3 for details. The clustering results are shown in Table 3. The results of the two system clusters pointed out that the five counties, Cuyahoga, Hamilton, Montgomery, Allegheny, And Philadelphia, are the two counties with the largest number of heroin cases and the most counties with the largest number of synthetic opioid cases. Counties are the top priority areas in the United States that require major concerns.

Table 3 Clustering result

Category	FIPS_Combined				
Category I	39035	39061	39113	42003	42101
Category II	other				

Time-Series Model

The panel data includes time series data and cross-section data. We have obtained five key counties (39061, 39035, 42101, 39113, 42003) in the previous section. Now we analyze the time and space characteristics of the data in these five counties. Based on the extracted features for clustering, the clustering results have depicted the absolute amount of specific drugs from this perspective. To extract more relevant information from the data provided by NFLIS, we use the county's specific drugs and the county's total The ratio of the number of drugs is derived data, and time series analysis is performed. The final results of the 2010-2017 consecutive year time series analysis of the heroine ratio and the synthetic opioid ratio of the five counties in the five counties are shown in the following Table 4.

Reference[5] mentions the statistical analysis of data from previous years, The number of cases of abusive use of opioids in a state accounted for about 20% of the number of cases of drug deaths in the state, which already reflects the seriousness of the abuse of opioids. We borrowed this ratio to indicate that when a county's abuse of opioids accounted for 20% of the number of drug use cases in the county, it indicated that the situation was critical and relevant government departments should pay great attention to this. Therefore, the arguments such as the position of the sequence indicating the "ratio" and the year when the ratio reaches 20% are brought into the dynamic panel data model, and the critical value of the independent variable Y_{it} that is, the threshold of the corresponding sequence can be obtained separately.

Table 4 Time series analysis results

Are there any	Variable	Model	Time-Series	R2	Are there any specific concerns in 2010-2027?	The earliest concerns in time 2010-2027	Threshold of
390	synthetic	Variable ARIMA(0,2,0) 0.905			YES	2017	----
35	opioids	ratio	Brown	0.909	YES	2018	6782.83
390	synthetic	Variable ARIMA(0,2,0) 0.889			YES	2016	----
61	opioids	ratio	Brown	0.893	YES	2017	5425
∶∶	∶∶	∶∶∶	∶∶∶	∶∶∶	∶	∶	∶
		Variable	Damped Trend Holt		YES YES		4985
421	heroin	ratio		0.98		2010	
1				0.869		2015	

Principal Component Evaluation Model Based on Entropy Weight Method

Data analysis and processing

We need to analyze the common variables of the counties shared by the seven-year data, and by comparison, we find that each observation has four forms (Estimate; Margin of Error; Percent; Percent Margin of Error), and what we need is the estimated amount, and the variables and counties in the annex from 2010 to 2016 are not exactly the same. The data of the four tables from 2010 to 2013 are the same as the individuals, and the variables of the three tables are the same as the individuals from 2014 to 2016. For example, from 2010 to 2013 (county code) GEO.id2 has 51515, and the county code from 2014 to 2016 does not have this county. From 2010 to 2013, there are no variables (COMPUTERS AND INTERNET USE - Total Households, COMPUTERS AND INTERNET USE - Total Households - With a computer, COMPUTERS AND INTERNET USE - Total Households - With a broadband Internet subscription). Therefore, the relevant individuals of 51159, 51161, 51685 in the variable GEO.id2 are deleted. In the same way, the variables common to the seven years are screened for analysis. After the above treatment, 149 variables and 464 samples (counties) were obtained. And we need to analyze the data of the first question. And similarly, screen out the same county in the first two questions, and finally get 460 samples. After excluding the variables with unreasonable estimates and error ranges, we will select the variables we need from the remaining variables according to the literature and topic requirements, and finally get 21 variables to classify the data such as educational achievements. Variables of less than nine years and education levels of nine to twelve years are aggregated, new indicators are obtained for education levels below 12 years, and so on, and the four key points included in the first question are included. From the variable to the second question indicator, the last 22 variables selected initially are shown in Appendix 4.

Grey correlation analysis

There are known information in the objective world, as well as many unknown and unconfirmed information. Known information is white, unknown or non-confirmed information is black, and between

the two is gray. The grey concept is the integration of the concepts of "less data" and "information uncertainty". The grey system theory is aimed at this kind of uncertainty problem with neither experience nor information, that is, the problem of "less data uncertainty". The grey system theory regards the uncertainty as the amount of gray. In essence, it is a mathematical theory to solve the uncertainty theory of information deficiency. Because the gray system has less data and incomplete information, it is difficult for decision makers to determine the quantitative relationship between factors. It is difficult to distinguish the main factors and secondary factors of the system, thus introducing the gray correlation analysis method. A comprehensive evaluation method based on the grey system theory--grey correlation analysis method is to measure the degree of correlation between factors according to the similarity or dissimilarity between the developmental trends of factors, and quantifies or orchestrate the factors between systems with incomplete information.

According to the theory of grey relational space, the original data needs to satisfy the dimensionless or the same dimension. In this paper, the extremum method is used to dimensionize the original data, and the processed data is combined with the ideal object data column to obtain a new matrix:

$$S = \begin{Bmatrix} 1 & 1 & 1 & 1 \\ S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \cdots & S_{mn} \end{Bmatrix}$$

Record $S_i = (S_{i1}, S_{i2}, \ldots_r S_{in}), i = 0, 1, \ldots m, S_0$ as the reference sequence, and calculate the correlation coefficient layer $\beta_i(i)$ of the jth index of S_i and the jth index of S_0 (i=1,2,…,m; j=1,2,…,n)

$$\beta_i(j) = \frac{\min\limits_i \min\limits_j |S_{oj} - S_{ij}| + \rho \max\limits_i \max\limits_j |S_{oj} - S_{ij}|}{|S_{oj} - S_{ij}| + \rho \max\limits_i \max\limits_j |S_{oj} - S_{ij}|}$$

In the above formula, p∈[0,1], generally takes p=0.5. By calculating as above, the correlation coefficient matrix β is obtained:

$$\beta = \begin{Bmatrix} \beta_1(1) & \beta_1(2) & \cdots & \beta(2) \\ \beta_2(1) & \beta_2(2) & \cdots & \beta(2) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_m(1) & \beta_m(2) & \cdots & \beta(2) \end{Bmatrix}$$

Let $x_i = \frac{i}{n} \sum\limits_{j=1}^{n} \beta_i(j)$, x_i is the degree of association between the i-th evaluated object and the ideal object. The merits of the object to be evaluated are evaluated according to the size of the x_i value. The larger the x_i, the higher the degree of association between the i-th evaluated object and the ideal object, and thus the better it is among all the evaluated objects.

Because the topic requires judging whether the use or use trend of opioids is related to the socio-economic data of the census, and the

part 2 data shows the rules of the first-level indicators and the second-level indicators, of which there are 7 first-level indicators. Combining the use and use trends of opioids (represented by the number of drug cases in each county and the number of drug cases in each county and the total number of identified drug cases), the gray correlation analysis is carried out on 9 indicators. The correlation degree is solved by using Matlab. The specific procedure is shown in Appendix 5. As can be seen from the above results, all correlations are greater than 0.5. As can also be seen from Appendix 3, the correlation coefficient matrix is close to 1, indicating that the use or use trend of opioids has a strong correlation with all aspects of the population (Appendix 6).

Principal component evaluation model based on entropy weight method

According to the idea of information entropy, entropy is an ideal scale when evaluating the index weight of indicator system. The principal component analysis method has a good dimensionality reduction processing technology, which can transform multiple indicators into several uncorrelated comprehensive factors, and the comprehensive factor variables can reflect most of the information of the original index variables, which can better solve many problems. Requirements for indicator evaluation. Therefore, a principal component evaluation model based on entropy weight method can be established.

Consider an indicator evaluation system, in which there are n evaluation indicators, m evaluated objects, and the raw data of the corresponding indicators of the evaluated objects are represented by the following matrix form.

$$R = \begin{Bmatrix} r_{11} & r_{12} & r_{1n} \\ r_{21} & r_{22} & r_{2n} \\ \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & r_{mn} \end{Bmatrix}$$

First, the raw data is dimensionless:

Remember that the optimal value for each column in R

$$r_j^* = \begin{cases} \max r_{ij}, where\ j = the\ yield\ index \\ \max r_{ij}, where\ j = the\ cost\ index \end{cases} \quad i=1,2,\ldots,m, j=1,2,\ldots n \quad (3.9)$$

(Note: The profitability indicator is that the larger the index value, the better. The cost index is the smaller the indicator value, the better.)

After the original data is dimensionless, it is recorded as a matrix $S = (s_{ij})_{m \times n}$

$$S_{ij} = \begin{cases} \dfrac{r_{ij}}{r_j^*}, where\ j = the\ yield\ index \\ \dfrac{r_{ij}}{r_j^*}, where\ j = the\ cost\ index \end{cases} \quad i=1,2,\ldots,m, j=1,2,\ldots n \quad (3.9)$$

Normalize S, remember

$$S_{ij}' = \frac{S_{ij}}{\sum\limits_j \sum\limits_i S_{ij}}$$

The $s'_{ij} \in [0,1]$ obtained in this way does not destroy the proportional relationship between the data.

Define the entropy of the jth evaluation indicator as

$$H_j = -k \sum_{i=1}^{m} t_{ij} \ln t_{ij}, j = 1, 2, \cdots n$$

$$t_{ij} = \frac{S'_{ij}}{\sum\limits_{i=1}^{m} S_{ij}}, j=1,2,....n, k=\frac{1}{1nm'}$$ (so the chosen k is such that,

$0 \le H_j \le 1$,convenient for subsequent processing)

Define the difference coefficient of the jth evaluation indicator as

$$\alpha_j = 1 - H_j, j = 1,2,\cdots n$$

Define the entropy weight of the jth evaluation indicator as

$$\omega_j = \frac{\alpha_j}{\sum\limits_{j=1}^{n}}, j = 1, 2, \cdots n \qquad (3.10)$$

The entropy weight thus defined has the following properties:

When the values of the evaluated objects on the index J are exactly the same, the entropy value reaches the maximum value of 1, and the entropy weight is zero, which means that the indicator does not provide any useful information to the decision maker, and the indicator can be considered to be cancelled.

When the values of the evaluated objects on the index J differ greatly, the entropy value is small and the entropy weight is large, which means that the indicator provides useful information to the decision maker, and in the problem, each object is in the There are obvious differences in indicators, which should be focused on;

The larger the entropy of the indicator, the smaller its entropy weight, and the less important the indicator is. The entropy defined by equation (3.10) satisfies:

$$0 \le \omega_j \le 1 \text{ and } \sum_{j=1}^{n} \omega_j = 1$$

It can be seen from the above discussion that the entropy weight method reflects the importance of the difference between the observations of the same indicator. The final indicators are as shown in Figure 7 below.
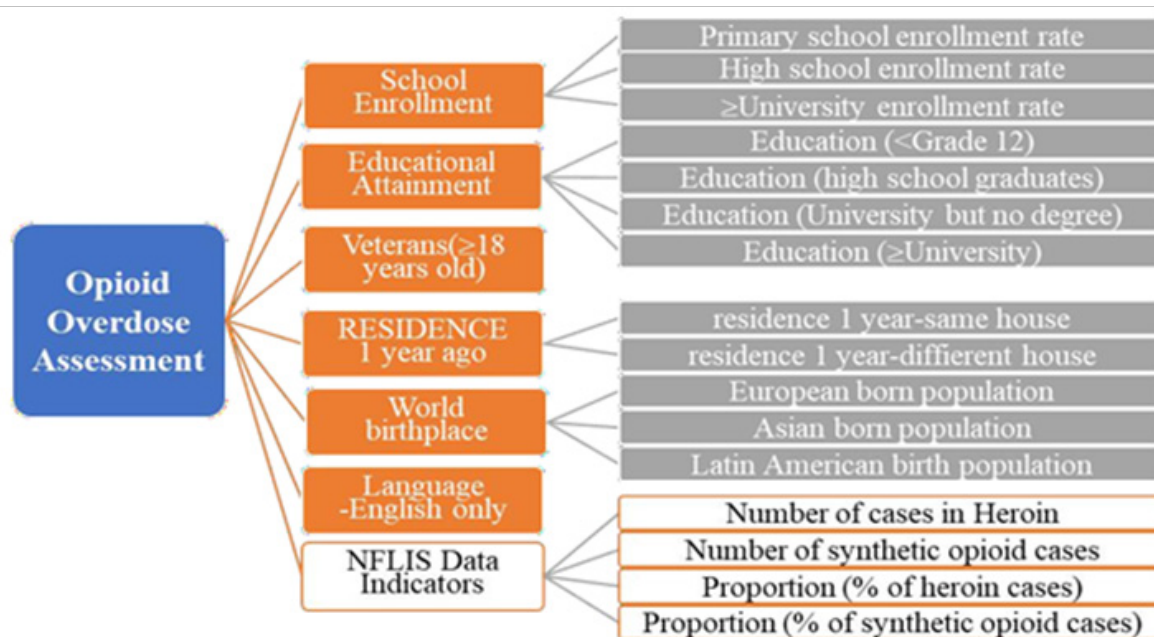

Figure 7 Opioid flooding indicator system.

Solution and result

The above 18 indicators can be obtained by SPSS factor analysis to obtain the factor load matrix and the variance interpretation ratio. The variance interpretation scale table is as Table 5, the first four components are extracted, and the factor load matrix is shown in Appendix 2. It can be seen from Table 6 that the first four principal components explain 85.115% of the overall properties, that is, 85% of the features can be explained according to the first four principal components. Therefore, the first four principal components are analyzed here.

The eigenvector of the principal component is

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix} = \begin{bmatrix} u_{11}\sqrt{\lambda_1} & u_{12}\sqrt{h} & u_{1m}\sqrt{\lambda_m} \\ u_{21}\sqrt{\lambda_1} & u_{22}\sqrt{\lambda_2} & u_{2m}\sqrt{\lambda_m} \\ u_{p1}\sqrt{\lambda_1} & u_{p2}\sqrt{h} & u_{pm}\sqrt{\lambda_m} \end{bmatrix}$$

Among them, u_{ij} represents the value of the factor load matrix, λ_i represents the eigenvalue, and a_{ij} represents the eigenvector. Therefore, the part of the corresponding feature vector is shown in Table 7, and the full content is shown in Appendix 2.

Table 5 Correlation

Correlation								
0.9219	0.9218	0.949	NaN	0.9383	0.9737	0.9825	0.9249	I

Table 6 Total Variance Explained

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
I	13.279	66.397	66.397	13.279	66.397	66.397
2	1.614	8.068	74.465	1.614	8.068	74.465
3	1.119	5.596	80.061	1.119	5.596	80.061
4	1.011	5.054	85.115	1.011	5.054	85.115
5	0.864	4.32	89.435			
⋮	⋮	⋮	⋮			
20	5.36E-05	0	100			

Table 7 Feature vector (part)

Component	I	2	3	4
GEO.id2	0.0148	-0.0527	0.4878	0.724
year	0.004	0.4794	0.4037	-0.2019
SCHOOL ENROLLMENT - Elementary school (grades 1-8)	0.2717	-0.0252	0.0009	-0.002
SCHOOL ENROLLMENT - High school (grades 9-12)	0.272	-0.0291	-0.0095	0.005
SCHOOL ENROLLMENT - College or graduate school	0.2654	-0.0307	-0.0113	0.009
EDUCATIONAL ATTAINMENT - Less than 12th grade, no diploma	0.2574	-0.0291	-0.0832	0.0249
EDUCATIONAL ATTAINMENT - High school graduate (includes equivalency)	0.2615	0.0449	-0.139	0.1233
EDUCATIONAL ATTAINMENT - Associate's degree	0.2637	0.0575	-0.0841	0.0786
WORLD REGION OF BIRTH OF FOREIGN BORN - Latin		...	...	...
America	0.2009	-0.211	0.4112	-0.2884
LANGUAGE SPOKEN AT HOME - English only	0.27	0.022	-0.0832	0.0676
Heroin	0.2179	0.2039	-0.2675	0.1243
synthetic opioid	0.1131	0.3928	-0.2912	-0.0358
Heroin ratio	0.0623	0.4353	0.2127	0.3481
synthetic opioid ratio	0.009	0.5282	0.1683	-0.3093

Therefore, the main components are:

$$F_k = \sum_{p}^{p} a_{kp} x_p, i = 1, 2, 3, 4$$

Among them, a_{kp} represents the corresponding feature vector of the k-th indicator in the i-th principal component, and x_p represents the p-th index.

In the expression of the first principal component, the coefficients of the 3rd, 5th, 6th, 7th, 8th, 9th, 10th, 11th, 12th, 13th, and 16th indicators are large, and the eleven indicators play a major role, so we can Think of the first principal component as a comprehensive indicator consisting of these eleven single indicators. The second,

third and fourth principal components are the same. Any event can be derived from the opioid flooding score as long as the indicator is known. For the purposes of this paper, the total score is weighted by the first four principal components, and the principal component pre-factors can be the respective variance contribution rates.

which is

$$F = 0.66397F_1 + 0.08068F_2 + 0.05596F_3 + 0.05054F_4$$

Opioid class of abuse level

According to the principal component evaluation model based on entropy weight method, the algebraic value F of the opioid drug flooding score is obtained.[6,7] According to the distribution of F value,

the degree of flooding of opioids is graded, and the filled area map and frequency distribution map are obtained. As shown in Figures 8 & 9. It can be seen from Figures 8 & 9 that the maximum F value is greater than 90,000, all the data falls within the interval [0, 1000000], and more are gathered in the interval [0, 450000].In order to effectively classify the degree of flooding, the interval [0, 400000] is subdivided, and the data of the degree of flooding between each cell is counted to obtain a frequency distribution map, as shown in Figure 10. It can be found that the comprehensive evaluation value of more than 3000 reports is in the interval [0, 450000], and there are almost no reports of more than 400,000. Only a few important events can be seen in Figure 8. For example, the F value of the county 42101 is 96, 7854.7 which is the highest, and opioids are the most rampant. In order to further discover the level of data, the data of this interval is refined according to the step-by-step refinement analysis method, and is divided into the figures as shown in Figure 10.
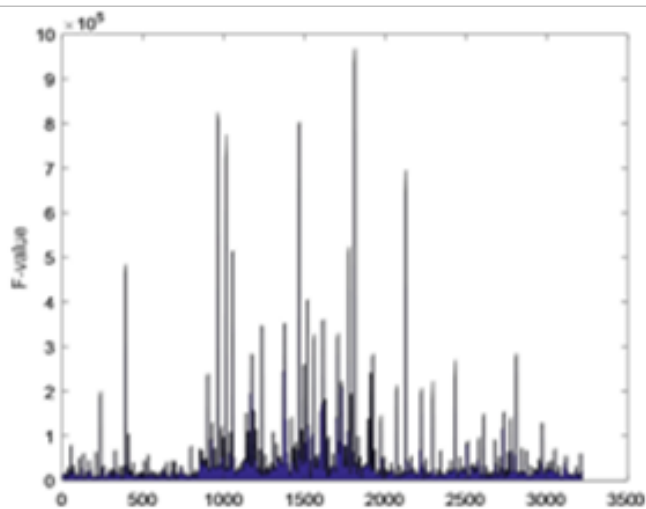
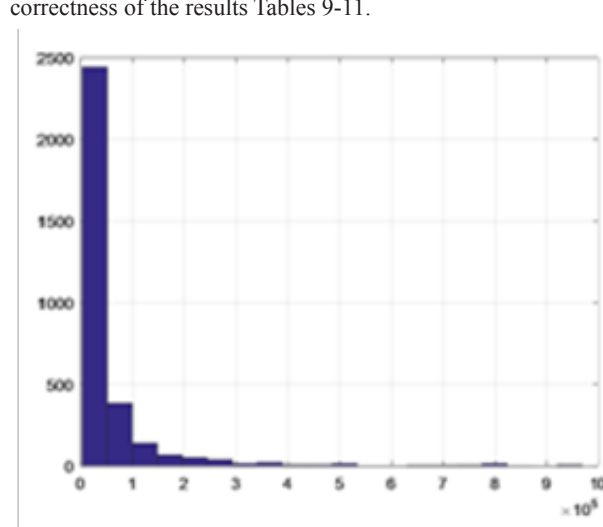From the Figure 10, we can see the obvious hierarchical distribution. As the interval is continuously refined, it is true that a large amount of data is found between [0,50000]. As the value of F is higher, the flood is more serious. Such incidents do occur in practice, and the use of appropriate opioids occurs in all regions of the United States, so the number is large. Therefore, according to the frequency of the degree of opioid influx, the degree of opioid influx is divided into one to three levels from high to low, see Table 8. The degree of flooding of opioids in Table 8 is graded, and level 1 indicates the greatest degree of flooding. In order to verify the rationality of the classification, combined with the actual considerations, random or selected boundary values for verification, the data can meet the requirements and meet the actual facts, which also proves the rationality and accuracy of the model. At the same time, the F values of the five counties (39035, 39061, 39113, 42003, 42101) that were firstly observed were in the serious category, further illustrating the correctness of the results Tables 9-11.

Figure 8 area.

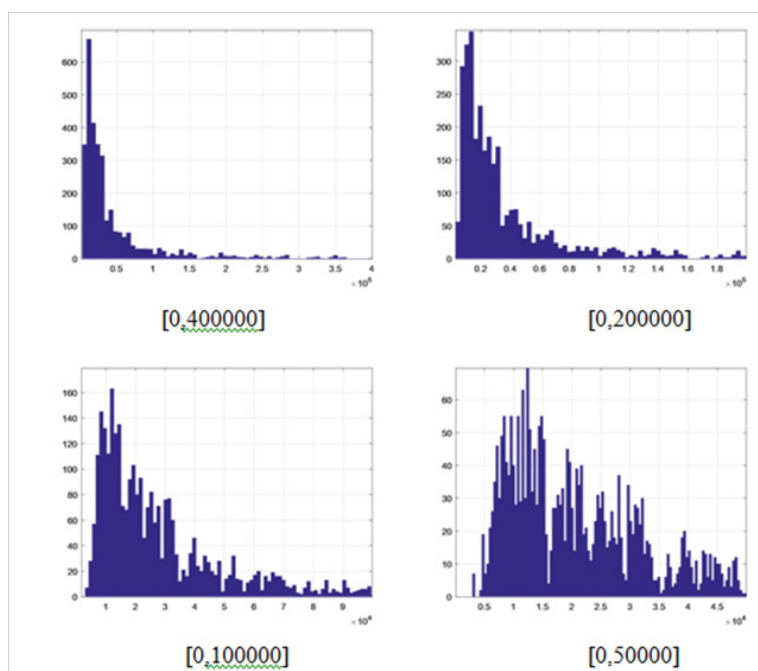**Figure 9** F-value frequency distribution.

Figure 10 Number of reports in different evaluation value intervals.

Table 8 Classification of the extent of opioids

Level	F-flooding score	Corresponding number of reports	Total ratio
1-serious	>400000	63	1.96%
2-General	(50000, 400000]	728	22.60%
3-lower	(0, 50000]	2430	75.44%

Table 9 Estimated interval

County code	Parameter a7 (high school education)	Parameter a8 (University but no degree)	Parameter a9 (≥university degree)	F1 value Corresponding interval	Flood level
39035	(0,0.2637)	(0.2615,1)	(0.2615,1)	(22524.912,41786.003)	Lower
39061	(0,0.2637)	(0.2615,1)	(0.2615,1)	(72094.505,334953,171)	General
39113	(0,0.2637)	(0.2615,1)	(0.2615,1)	(28967.624,48021.552)	Lower
42003	(0,0.2637)	(0.2615,1)	(0.2615,1)	(26349.411,29854.518)	Lower
42101	(0,0.2637)	(0.2615,1)	(0.2615,1)	(21552.004,45897.343)	Lower

Table 10 Forecast data 1

Time	Predicted value of synthetic opioids	V39061 Predicted value of synthetic opioids	V42003 Predicted value of synthetic opioids	V39035 Predicted value of synthetic opioids	V42101 Heroin case number prediction	V39061 Heroin case Number prediction
2010	23.83	986.33	392	224.57	3360.48	2044.22
2011	200.83	760.33	301	973.2	3575.35	2357.88
2012	380.83	478.33	239	1035.47	3745.12	2675.21
2013	419.83	451.33	251	2242.11	3921.44	3146.27
2014	411.83	500.33	280	2916.63	4074.26	3552.92
2015	661.83	851.33	386	3262.71	4193.2	3921.72
2016	1050.83	3385.33	1003	3139.86	4470.1	4314.29
2017	3876.83	4528.33	1705	2839.98	4745.26	4699.54
2018	6782.83	8444.33	3757	2440.05	4971.38	4885.61
2019	9750.5	11900	5479	2246.33	5210.89	5226.22
2020	13090	15792	7419	2074.26	5461.93	5566.84

Table 11 Forecast data 2

time	Prediction _of_ the _proportion _of_ synthetic _opioids39035	Prediction _of_ the _proportion _of_ synthetic _opioids39061	Prediction _of_ the _proportion _of_ synthetic _opioids39113	Prediction _of_ the_ ratio _of _heroin39035	Prediction _of_ the _ratio_ of _ heroin42101
2010	0.006468	0.007762	0.008335	0.140011	0.107245
2011	0.004262	0.007403	0.008581	0.140069	0.129526
2012	0.002056	0.006179	0.008827	0.141602	0.149719
2013	0.000667	0.002216	0.004667	0.138233	0.168613
2014	0.002581	0.004688	0.014995	0.139344	0.187054
2015	0.017201	0.029276	0.08438	0.142694	0.207191
2016	0.045115	0.178035	0.046437	0.147687	0.233344
2017	0.197334	0.222675	0.226049	0.150012	0.254845
2018	0.265979	0.320816	0.465621	0.146909	0.270898
2019	0.342783	0.396373	0.630162	0.147922	0.290963

Table Continued....

time	Prediction _of_ the _proportion _of_ synthetic _opioids39035	Prediction _of_ the _proportion _of_ synthetic _opioids39061	Prediction _of_ the _proportion _of_ synthetic _opioids39113	Prediction _of_ the _ratio _of _heroin39035	Prediction _of_ the _ratio_ of _ heroin42101
2020	0.419587	0.47193	0.794704	0.148945	0.311027
2021	0.496391	0.547488	0.959245	0.149976	0.331091
2022	0.573196	0.623045	1.123786	0.151017	0.351155
2023	0.65	0.698602	1.288327	0.152066	0.371219
2024	0.726804	0.77416	1.452868	0.153125	0.391283
2025	0.803608	0.849717	1.617409	0.154194	0.411347
2026	0.880412	0.925275	1.78195	0.155272	0.431411
2027	0.957217	1.000832	1.946491	0.15636	0.451475

Linear programming model

The foundation of model

In question 1, we have used system cluster analysis and time series analysis to select the five counties where opium is the most widespread in the United States. In question 2, we sorted the weighted composite scores F of 460 counties, and divided 460 counties into three layers according to the order of F. The larger the value of F, the higher the extent of opioids in the county. Regarding question 3, we find that if min F is regarded as an objective function, a linear relationship can be established between the socio-economic secondary indicators used and their own primary indicators. That is, all 20 indicators used can be formed into restrictions. Further, since F1's contribution rate is 66%, most of the information about the whole can be explained. Considering the implementation cost of the strategy, in order to make the effectiveness of the anti-opioid crisis strategy as obvious as possible, we replace min F with min F1. The linear programming model is established as follows.

$$\text{Objective function: } \min F_1 = \sum_{k=1}^{20} a_k \chi_k$$

$$s.t. \begin{cases} x_1 = GEOid\,2, x_2 = year, x_3 + x_4 + x_5 \leq enrollment\,rate, \\ x_6 = 0, x_7 \geq population - Education\,(high\,school\,graduates), \\ x_8 \geq Population\,Education\,(University\,but\,no\,degree), \\ x_9 \geq Population\,Education\,(\geq University), x_6 + x_7 + x_8 + x_9 = Population\,Education, \\ x_{10} = Veterans\,(\geq 18\,years\,old), x_{13} + x_{14} + x_{15} \leq world\,born\,population, \\ x_{11} + x_{12} \leq The\,number\,of\,people\,living\,in\,a\,house\,for\,one\,y\,ear, \\ x_{16} \leq Family\,language\,population, x_{17} + x_{18} \leq number\,of\,opioid\,cases\,per\,county, \\ x_{19} + x_{20} \leq proportion - number\,of\,opioid\,cases\,per\,county\,total\,county, \\ x_1, x_2, x_3, ... x_{18}, x_{19}, x_{20} \geq 0, \end{cases}$$

Among them, X_k is the 20 indicators selected by Philadelphia, and a_k is the corresponding coefficient (parameter).

Model solution and sensitivity analysis

Therefore, we can use the adjustment of an indicator X_k in the constraint as a strategy against the opioid crisis. And through the local sensitivity analysis after the change of X_k, the parameter range (c, k) of each index is obtained. That is, when the parameter of an indicator is in (c, k), the optimal solution does not change. It can be preliminarily understood that when the parameters of an indicator fluctuate in (c, k), the provided strategy is effective. Of course, we can also bring the obtained parameter range into the first principal component score expression to find a new F1 value, and determine whether the parameter range is valid according to the level of the opioid drug flood level to which the new F1 value belongs.

When the new F1 value is at the third level, the parameter range is successful. When it is at the second level, the success or failure of the parameter range is not obvious, that is, the corresponding measures are not effective. When at the third level, this parameter range is unsuccessful. Taking the indicator of education as an example, we give measures to resist the opioid crisis: education is a comprehensive indicator of importance. Our strategy is to continue to increase the popularity of basic education in the United States, so that all people over the age of 25 in the United States will reach at least the high school education and above. The more a person knows about the dangers of drugs, the more proficient the correct use of opioids, the higher the knowledge and personal cultivation, the less likely he is to take drugs and abuse opioids. Combined with the above analysis, in the new constraints, the population below the high school education is 0, and the reduced population is distributed to the population with higher education. As we analyze the latest data provided, the data of the five counties of Cuyahoga, Hamilton, Montgomery, Allegheny, And Philadelphia are substituted into the model. Analyzing the sensitivity of important coefficients with Lingo and the estimated range of the parameters can be obtained. Substituting the endpoint value of each parameter into the first principal component score, the interval of the opioid flooding scores of the five counties was obtained, and the

results are shown in the following table. According to Question 2, the grading model of the opioid flooding scores in the five counties shows that by raising the basic education level of young people under the age of 25, only Hamilton County, Ohio (39061) is in the general level of flooding, and the remaining counties have fallen to lower levels. The level of opioids in these five counties has dropped from severe to lower or general, indicating that the future opioid crisis predicted in Part 1 may not occur. From the overall situation of the five counties, our strategy is effective.

Model evaluation and promotion

Strengths

In the second problem, the panel data is used in the dynamic panel data model. Compared with the cross-section data model, the panel data model controls the deviation of the OLS estimation caused by the unobservable variables, making the model more reasonable and the sample estimation of the model parameters more accurate. Compared with time series data, the panel data model expands the sample information, reduces the collinearity between variables, and improves the validity of the estimator. Among them, the dynamic panel data model can more accurately adjust the dynamics of the response variables. In question 2, the degree of flooding of opioids was based on the frequency of F values, and the results were verified.

Weaknesses

The data given in question 2 does not take into account income, economic indicators.

In question 2, the degree of spread of opioids is divided into three layers, with certain subjectivity.

Discussion

Since some research results are not the focus of answering the question, and considering the reasons for the paper, the data we have obtained in the modeling process are not all in the text. However, some models have good statistical results, so we want to put the data passed by the statistical test in the memorandum. The following table shows the data prediction results for a time series analysis during the modeling process. It should be pointed out that since the known data is only 8 years, we believe that the reliability of the data in the later years may be verified. And our study is limited to the extent that it focuses on the data provided by NFLIS concerned with opioid crisis in the US at one time period (2010–2017). After sorting and analyzing the panel data, we decided to transform the derived data and then model the panel data, cross-section data and time series data respectively.[9,10] Next, we consider how to objectively select a large number of socio-economic data and indicators, and then establish a model that can reflect two different databases at the same time, so that the model can be combined with some indicators. Further, we note that the sample selection strategy may have resulted in an underrepresentation of heroin users with a prescription opioid misuse history. Additionally, we note that the findings reported here may not be completely generalizable to other settings and time periods.[11,12]

Acknowledgments

Conflicts of interest

We have no conflict of interests to disclose and the manuscript has been read and approved by all named authors.

References

1. Yang Yuhui, Xu Xiuli, Zhu Zhu. Overview of Opioid Abuse and Its Governance in the United States. *Chinese Medical Alert*. 2017;14(12):746–751.

2. Wang Shujiang, Li Weiyan. Dose titration and conversion of opioids in the treatment of cancer pain. *Journal of Southeastern Defense Medicine*. 2016;18(5):522–526.

3. Yan Shuai, Dang Yaoguo, Ding Song, Shang Zhongju. Research on panel data clustering method based on grey likelihood function. *Control and decision*. 2019;1–7.

4. Wu Jian, Liu Zidong, Wang Chao, et al. Analysis of Resonance Characteristics of Inverter Grid–Connected System Based on Sensitivity Theory. *Journal of Electric Machines and Control*. 2018;22(12):11–29.

5. Overdose Death Rates. 2018.

6. Volkow ND, Collins FS. The role of science in addressing the opioid crisis. *N Engl J Med*. 2017;377(4):391–394.

7. Zhao Hailong. Research and application of function substitution method for reliability and reliability sensitivity analysis. *Northwestern Polytechnical University*. 2015.

8. Xu Chongang, Hu Yuanman, Chang Wei, et al. Sensitivity analysis of ecological model. *Chinese Journal of Applied Ecology*. 2004;15(6):1056–1062.

9. Zhu Xiaohua, Yang Xiuchun. Application of Analytic Hierarchy Process in Regional Eco–environmental Quality Assessment. *Scientific and Technological Management of Land and Resources*. 2001;5:43–46.

10. Li Xueping. Discussion on the Method of Scoring Index Weights by Analytic Hierarchy Process. *Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition)*. 2001;1:25–27.

11. Dang Yaoguo, Hou Yuqing. Multi-indicator panel data clustering method based on feature extraction. *Statistics & Decision*. 2016(19): 68–72.

12. Cicero TJ, Ellis MS, Kasper ZA. Increases in self-reported fentanyl use among a population entering drug treatment: The need for systematic surveillance of illicitly manufactured opioids. *Drug and Alcohol dependence*. 2017(177):101–103.

13. Dang Yaoguo, Hou Yuqing. Multi–indicator panel data clustering method based on feature extraction. Statistics & Decision, 2016(19): 68–72.

14. Cicero, T. J., Ellis, M. S., Kasper, Z. A. Increases in self–reported fentanyl use among a population entering drug treatment: The need for systematic surveillance of illicitly manufactured opioids. Drug and Alcohol Dependence, 2017(177): 101–103.