

Received February 24, 2020, accepted March 6, 2020, date of publication March 10, 2020, date of current version March 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2979859

Study on Prediction Model of HIV Incidence Based on GRU Neural Network Optimized by MHP SO

XIAOMING LI^{1,2}, XIANGHUI XU³, JIE WANG¹, JING LI¹, SHENG QIN¹, AND JUXIANG YUAN¹

¹School of Public Health, North China University of Science and Technology, Tangshan 063210, China

²Hebei Province Key Laboratory of Occupational Health and Safety for Coal Industry, North China University of Science and Technology, Tangshan 063210, China

³Department of Internal Medicine, North China University of Science and Technology Hospital, Tangshan 063210, China

Corresponding author: Juxiang Yuan (gwxyjxb@ncst.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFC0900605, and in part by the Science and Technology Research Program of Colleges and Universities in Hebei Province under Grant QN2019190.

ABSTRACT Acquired Immune Deficiency Syndrome (AIDS) is still one of the most life-threatening diseases in the world. Moreover, new infections are still potentially increasing. This difficult problem must be solved. Early warning is the most effective way to solve this problem. Here, we aim to determine the best performing model to track the epidemic of AIDS, which will provide a methodological basis for testing the time characteristics of the disease. From January 2004 to January 2018, we built four computing methods based on AIDS dataset: BPNN model, RNN model, LSTM model and MHP SO-GRU model. Compare the final estimated performance to determine the preferred method. Result. Considering the root mean square error (RMSE), mean absolute error (MAE), mean error rate (MER) and mean absolute percentage error (MAPE) in the simulation and prediction subsets, the MHP SO-GRU model is determined as the best performance technology. Estimates for the period from May 2018 to December 2020 suggest that the event appears to continue to increase and remain high.

INDEX TERMS AIDS, deep learning, incidence prediction, LSTM network, MHP SO-GRU network, RNN.

I. INTRODUCTION

Acquired immune deficiency syndrome (AIDS) is the result of human immunodeficiency virus (HIV) infection, and its main feature is T cell immune function defect [1]. HIV mainly exists in the blood, semen, vaginal secretions and milk of infected people. Therefore, it is usually transmitted between people through direct sexual contact, iatrogenic and perinatal transmission [2], [3]. Since 1997, the annual incidence of AIDS in the world has declined slightly, but it is still one of the most important causes of death in the world's major disease burden [4]–[7]. It is estimated that there were 1.8 million new infections in 2016, with an average life expectancy of 40.32 years. About 70% of these new infections occur in low and middle-income countries, with investment spending reaching 8.13 billion US dollars [8]. In recent years, the number of AIDS cases reported annually

has increased [9], [10]. In addition, the mortality rate of AIDS in class B statutory infectious diseases is the highest from 2008 to 2017, and the vaccine has not been determined [11], [12]. Therefore, facing the major public health challenges and threats brought by AIDS, and providing a clear and quantitative direction for effective formulation of preventive plans and rational allocation of limited available resources, this is a high-precision and accurate understanding of the prediction model of AIDS epidemic behavior is necessary. Jia J established an HIV / AIDS epidemic model with general nonlinear incidence and treatment methods, and studied the stability of disease-free balance and unique local balance [13]. Wang Yawen established the combination of autoregressive moving average model, generalized regression neural network and ARIMA model. ARIMA (1,1,1) (0,1,1) 12 and ARIMA-GRNN model can better predict the monthly incidence of AIDS in China, but the effect of the combined model is better [14]. Yang K used nonparametric local smoothing method to develop a flexible and effective

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei¹.

data modeling method for St disease incidence, monitoring the incidence of cancer, AIDS, cardiovascular disease and other chronic or infectious diseases, with good results [15]. Mar Masi áused the multi state method to predict the development of PLWH into non AIDS events, and estimated three transfer probabilities: from alive and NAE free to alive and NAE experienced (“NAE development”); from alive and NAE experienced to death (“death after NAE”) [16]. Rezaei s elucidates the overall prevalence of depression in HIV / AIDS patients through systematic review and meta-analysis [17]. Adamu incidence rate in various states affected by insurgency activities in Boko Haram, Nígeria, collected data from 2004 to 2017, reporting 16 [18]. The information of 102 patients, including age, gender, diagnosis year and patient status, found that most of the HIV / AIDS infected people were women in these Boko Haram areas. Although these laws were discovered, they did not give a trend of future incidence rate. Deribew A studied the development trend of HIV / AIDS mortality in Ethiopia from 1990 to 2016, found that the decline of HIV / AIDS mortality was relatively slow, and provided the corresponding HIV testing and treatment plan, but did not give a future development trend, and the plan had a certain lag [19].

Deep learning is a kind of machine learning method to represent data washing. It combines the bottom features of data into the top features through deep neural network, and is used for classification, prediction and other tasks. Compared with the traditional shallow learning methods such as neural network, logic regression, support vector machine, decision tree, etc., its network depth is deeper, its nonlinear representation ability is stronger, and its network learning ability is stronger. Shallow learning is often ineffective in dealing with complex problems, even when the number of samples is sufficient. This is because the network complexity is low and the learning ability is not strong. Even if a large number of samples are input, some important details cannot be learned. However, in the past, due to the limitation of computing power, blindly increasing the number of layers and parameters of the network is easy to lead to slow convergence and over fitting of the network. Today’s deep learning technology can effectively solve these problems. While increasing the network depth, greedy layer by layer training method is used to prevent over fitting and improve the network performance.

II. RNN AND ITS DERIVED NEURAL NETWORK

From the first Turing test to the present deep learning, in recent decades, imitating human brain has always been the direction of artificial intelligence. However, in the past few years, due to the limitation of computing power, computers have been unable to process the same amount of data as human brain. With the rapid development of computer computing power, deep learning model solves this problem through various methods [20]. ANN (artificial neural network) has also developed rapidly and been widely used, and derived convolutional neural network (CNN) with spatial distribution data and recurrent neural network (RNN, recurrent

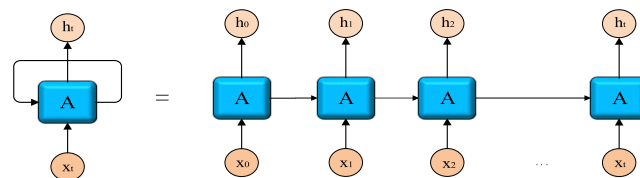


FIGURE 1. RNN cycle expansion.

neural network) with temporal distribution data [21]. At the same time, the cyclic neural network model is widely used in time series data processing because it can process the time-series data.

A. CYCLIC NEURAL NETWORK MODEL (RNN)

In terms of time series data processing, taking the prediction of AIDS incidence to be discussed in this paper as an example, it is different from the traditional artificial neural network (ANN) that takes a certain time cross-section factor as the input value of data. The RNN takes a certain long-term factor data as the time series and takes the data in the past period as the input value [22].

Figure 1 is to expand unified neuron A according to time [0, t]. the characteristics of chain structure reveal that RNN is essentially related to sequence and list. In this process, we can see the operation of the whole neuron at different time points. The input variable x0 is processed in the neuron at the time of 0 and input h0. This process is not over, the whole process has entered into 1 moment.

However, due to the high complexity of RNN model, gradient explosion and gradient disappearance are easy to occur when the input variables enter the hidden layer for data processing. In 1994, yoshua bengio explained the gradient disappearance by mathematical formula (1) [23].

$$h_t = \delta (Ux_t + Wh_{t-1} + b) \tag{1}$$

Because the hidden layer gradient will multiply the coefficient w repeatedly in the process of back propagation along the time dimension, if the data is long series data, the more times the coefficient W is multiplied repeatedly, the more likely it is to cause the gradient explosion and gradient disappearance.

B. LONG AND SHORT TERM MEMORY NETWORK MODEL (LSTM)

The solution to RNN problem comes from the introduction of long-term memory network (LSTM) in 1997. Hochrater and schmiduber, the proponents of long-term and short-term memory networks, found that the essential reason for RNN to produce gradient explosion and gradient disappearance is that the calculation method of hidden layer state causes the gradient to be expressed in the form of continuous product, and put forward an improved way to solve this problem [24], as shown in Figure 2.

If we omit ot, yt and Lt for each layer, then the model can be simplified as shown in the figure above. In RNN model,

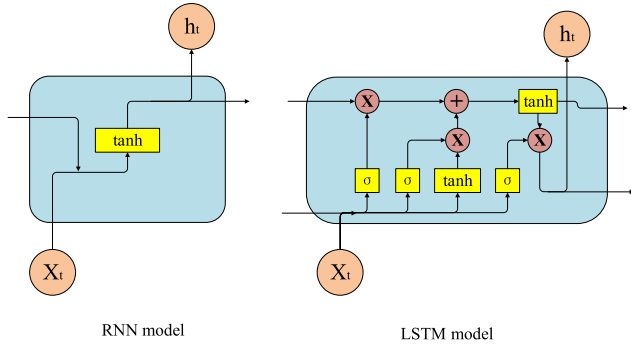


FIGURE 2. Internal schematic diagram of RNN model and LSTM model.

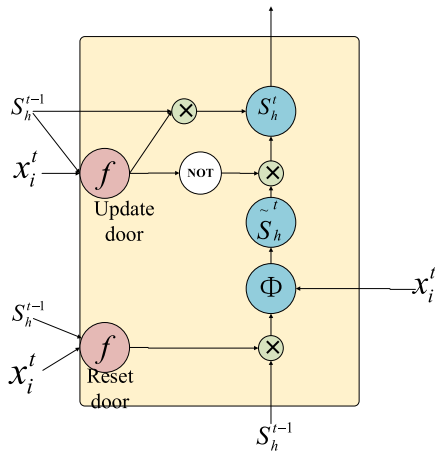


FIGURE 3. GRU unit structure diagram.

we can see clearly that the hidden state h_t is determined by h_{t-1} and x_t together. On the way through the activation function σ , which is \tanh in general, the LSTM model on the right side of the figure above is much more complex.

RNN neural network in the process of transforming the growing neural network, because of repeated cycles in the hidden layer, the error will become more significant along with the cycle, resulting in the disappearance of gradient, the change of training results is not obvious, and finally the whole training process cannot escape from the local optimal solution [25], [26].

III. GRU NEURAL NETWORK

GRU network was proposed in 2014, which is essentially a cyclic neural network and an improved version based on LSTM network [27]. Based on the LSTM unit, the input gate and forgetting gate in LSTM are combined into a single update gate, and the specific structure is shown in Figure 3.

$$a_u^t = \sum_{i=1}^I w_{iu}x_i^t + \sum_{h=1}^H w_{hu}s_h^{t-1} \quad (2)$$

$$s_u^t = f(a_u^t) \quad (3)$$

$$a_r^t = \sum_{i=1}^I w_{ir}x_i^t + \sum_{h=1}^H w_{hr}s_h^{t-1} \quad (4)$$

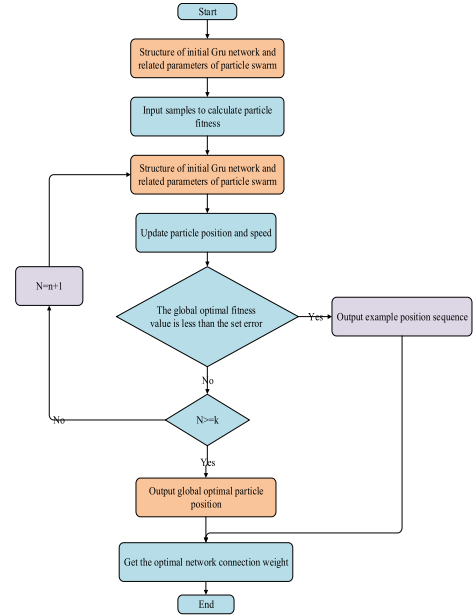


FIGURE 4. Flow chart of MHPSO-GRU model.

$$s_r^t = f(a_r^t) \quad (5)$$

$$\tilde{a}_{h'}^t = \sum_{i=1}^I w_{ih'}x_i^t + s_r^t \sum_{h=1}^H w_{hh'}s_h^{t-1} \quad (6)$$

According to Figure 3, the forward propagation formula can be obtained as shown in formula (2)-(6). Where u is the index of the updated gate vector, r is the index of the reset gate vector, h is the index of the hidden cell vector at all times, h' is the index of the new memory cell at t , and f and φ are activation functions. f is the sigmoid function, φ is the \tanh function, and $\tilde{s}_{h'}^t$ represents the information of the new memory unit at time t [28]. From Figure 4, it can be seen that the new memory information is obtained from the synthesis of the hidden unit states at t time input: x_i^t and $t - 1$ time, indicating that the new memory can have the new input information and historical information at the same time. The reset gate determines the weight of the hidden unit information in time $t - 1$ to the new memory information $\tilde{s}_{h'}^t$ in time [29]. If the current time state is independent of the previous time, the reset door can prevent the transmission of the previous time information. However, there is an update gate as the final output control gate, and the information of the previous moment can also be transmitted through the update gate. The weights of the new memory information and the hidden unit information at time: $t - 1$ are respectively controlled by the update gate. The above is the basic working principle of GRU, which can not only solve the problem of RNN gradient disappearing, but also reduce the network parameters compared with LSTM, and has faster convergence speed. Moreover, experiments show that the network performance is even better than LSTM.

A. GRU NEURAL NETWORK BASED ON MHPSO OPTIMIZATION

Particle swarm optimization (PSO) is a kind of swarm intelligence optimization algorithm, which is proposed by Kennedy and Eberhart in 1995 to simulate the foraging of birds. According to formula (7) and (8), the particles of the population constantly adjust their speed and position to optimize until the convergence termination condition is met.

$$V_{i,j}^{t+1} = \omega V_{i,j}^t + c_1 r_1 (y_{i,j}^t - x_{i,j}^t) + c_2 r_2 (y_i^t - x_{i,j}^t) \quad (7)$$

$$X_{i,j}^{t+1} = X_{i,j}^t + V_{i,j}^{t+1} \quad (8)$$

In the formula, $V_{i,j}^t$ represents the velocity of the particle i in the j -th dimension at the t -th iteration; $x_{i,j}^t$ is the position of the particle i ; ω is the inertia weight; c_1 and c_2 are acceleration coefficients or learning factors; $y_{i,j}^t$ is the t Individual extreme points during the iteration; y_i^t is the global extreme point of the particle swarm; r_1, r_2 are random numbers uniformly distributed in the interval $[0,1]$; $V_{i,j}^t \in [-V_{\max}, V_{\max}]$, V_{\max} are constants.

Aiming at the problem that particle swarm algorithm is easy to fall into the local optimal position, this paper proposes a hierarchical heterogeneous dynamic particle swarm optimization algorithm. The main purpose of this algorithm is to establish a direct connection between the particle swarm structure layer and the layer. The number of particles is the same.

Each time the algorithm iterates, it arranges all particles according to the current fitness value and assigns them to different levels. The adaptive value of the particles is inversely proportional to the level they are in. When the algorithm runs, the particles will be attracted by the particles on the upper layer, these particles are called attracted particles, and the particles themselves will also act as attracted particles of a certain particle in the lower layer. Particles, and it is the attracting particles of the fourth layer of particles; for the particles at the highest and lowest layers, its attracting particles are other particles in the same layer as it. In the process of optimization, the particles will not only move to the local best and global best positions, while also moving in the direction determined by its attracting particles.

The MHPSO algorithm incorporates a parameter related to attracting particles in the iterative update formula. The specific display can be seen from the following equation:

$$V_{i,j}^{t+1} = \omega V_{i,j}^t + c_1 r_1 (y_i^t - x_{i,j}^t) + c_2 r_2 (y_i^t - x_{i,j}^t) + \sum_{a=1}^{A_j} c_3 r_3 (x(i)_{a,j}^t - x_{i,j}^t) \quad (9)$$

Among them, $x(i)_{a,j}^t$ represents the position of the attracting particle a of the particle i , and B represents the attracting particle a of the particle i , c_3 represents the constant of the acceleration coefficient, and r_3 represents the attraction coefficient of particle i to attract particle a .

The position update formula is shown in equation (10):

$$X_{i,j}^{t+1} = X_{i,j}^t + V_{i,j}^{t+1} \quad (10)$$

The gradient descent method adopted by GRU neural network is easy to make the weight converge to a certain value, but it can not guarantee that it is the global minimum value of the error plane. The optimization of GRU neural network by MHPSO is to replace the gradient descent method in GRU neural network by GRU. The main idea of this algorithm is to optimize the connection weight and threshold of GRU neural network by MHPSO algorithm, which can improve the GRU neural network Prediction performance.

The specific steps of short-term traffic flow prediction based on optimized GRU neural network parameters of MHPSO are as follows:

(1) build training data set and test data set, and preprocess traffic data to obtain time series of historical traffic flow:

(2) using MHPSO to optimize the parameters of GRU neural network to get the optimal short-term traffic flow prediction model. The specific modeling steps are as follows:

Step 1: initialize the parameters of particle swarm and GRU neural network. The parameters of particle swarm include population scale, population layers, iterations, learning factors, and the limited range of particle position and velocity. The initial values of particle position and velocity are random. The parameters of GRU neural network mainly include the number of neurons in each layer of neural network and the number of layers of hidden layer.

Step 2: calculate the fitness value of each particle in the population and build the population hierarchy. The fitness function of particles in the population is defined as:

$$fit_i = \frac{1}{n} \sum (Y_i - y_i)^2, \quad i = 1, 2, 3, \dots, n \quad (11)$$

where n is the population size, Y_i is the sample output value, and y_i is the actual output value

Step3: calculate the fitness value of each particle

Step 4: update the speed and position of particles according to formula (3) and (4)

Step5: if the end condition of iteration (good enough position or maximum number of iterations) is reached, it will end. Otherwise, turn to step 2 to continue iteration MHPSO-GRU network is to optimize the initial weight of deep GRU network through MHPSO, and then restore the optimal solution as the initial weight of deep GRU network. Because the optimal solution obtained by MHPSO is close to the final required weight and width of deep GRU network in the global scope, the deep GRU network trained by learning can have less learning time and improve its disadvantage of easily falling into local minimum, thus improving the performance of prediction using this model.

B. PREDICTION MODEL OF AIDS INCIDENCE BASED ON GRU NETWORK

This paper proposes a prediction model of HIV incidence based on GRU network. The network structure of the model is shown in Figure 5. Among them, the one month incidence of AIDS as a sample, recorded as P . P is the incidence data of six months before the forecast month. The reason for setting

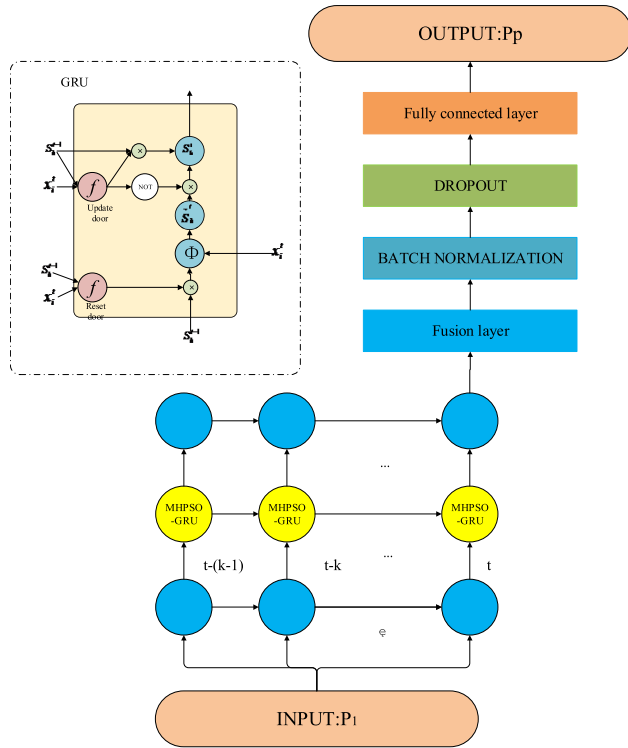


FIGURE 5. Prediction model of AIDS incidence based on MHP SO-GRU network.

the forecast interval to 6 months is based on the influence of date factors. As the input of RNN, P needs to be divided into multiple subsequences, and the time step of segmentation takes the middle value.

IV. MHP SO-GRU NETWORK AND PREDICTION OF AIDS INCIDENCE

A. GRU NEURAL NETWORK BASED ON MHP SO OPTIMIZATION

With the development of machine learning algorithm and data mining technology, the accuracy of incidence prediction is also increasing. In the face of complex and random systems, we often use one-dimensional time series data analysis, but this result is caused by many factors. It is necessary to study the influence of other factors in time series, which makes the pattern recognition of time series more important. In order to explore how these influencing factors affect the system and its internal relations, we can decompose the time series into some interpretable contents, such as trend item, cycle item, cycle component and irregular component, which can help us better simulate the time series. It is necessary to construct a topological space equivalent to the original system, that is, to project the influence factors hidden in the one-dimensional system that have not been observed into the geometric space, from the current state of the system, we can get the state of the next moment, that is, from the value at this time to the prediction value of the next moment of the time series. In this paper, GRU, a derivative network of RNN neural network, is used to predict the incidence of AIDS. The MHP SO-GRU

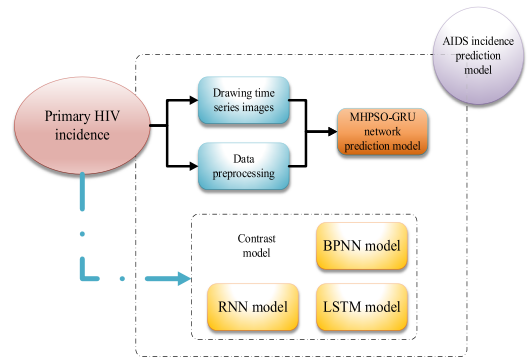


FIGURE 6. Prediction framework of AIDS incidence based on MHP SO-GRU model.

network prediction model was constructed for the sequence of AIDS incidence.

B. PREDICTION MODEL OF AIDS BASED ON MHP SO-GRU

Cyclic neural network can not only predict the image of time series, but also process the length of any input or output. The parameters of MHP SO-GRU model are less than those of traditional RNN and LSTM model, so it is easier to program and put into use. The incidence of AIDS over the years is a time series, which is more suitable for RNN model training. In this paper, we build a prediction model of AIDS incidence based on MHP SO-GRU model, and the specific framework is shown in Figure 6.

In the prediction framework of AIDS incidence shown in Figure 6, this paper first collects the time series data of AIDS incidence in 2004-2011, draws the time series image, and analyzes the trend impact of AIDS incidence. In this paper, a model based on MHP SO-GRU neural network is established to predict the incidence of AIDS in the next few years. At the same time, we also use BPNN model, traditional RNN model and LSTM model to model and predict the future incidence of AIDS based on historical data, and analyze the differences and advantages and disadvantages of the four prediction methods.

V. MODEL PERFORMANCE EVALUATION EXPERIMENT AND ANALYSIS

A. EXPERIMENTAL DATA AND NETWORK PARAMETER SETTING

The monthly cases of longitudinal AIDS collected from January 1, 2004 to April 30, 2018 in mainland China were collected from the National Notifiable Infectious Diseases Reporting System (2018). A total of 172 observations over 15 years were collated and summarized (table S1). Subsequently, in order to evaluate the maintainability and flexibility of each model developed, we used the data from January 1, 2004 to December 31, 2016 to train the basic parameters of the model, and used the remaining 16 data points to verify the model extrapolation ability. Our current research is not necessarily ethically approved or agreed, because the data do

TABLE 1. Predicted network structure.

Network layer	Cell number
Input : p_1	(1,6)
GRU	300
Fusion layer	400
BATCH NORMALIZATION	(1,6)
DROPOUT	400
Fully connected layer	100
Output: p_p	6

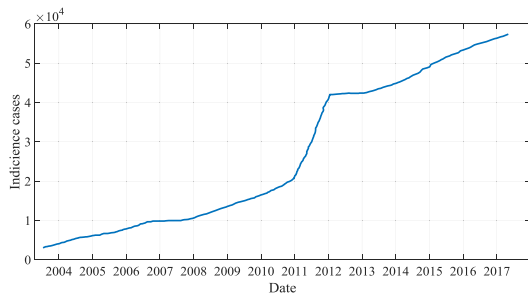


FIGURE 7. Time series of monthly AIDS cases from 2004 to 2017.

not have personally identifiable information and are publicly available.

B. TEST RESULTS AND COMPARISON

Between January 2004 and April 2018, there were 458078 AIDS cases in total, with an average of 2664 cases per month, 143 standard errors, and an average annual incidence of 2 per 100000 people. In 2017, China reported 57194 HIV cases per month, 18 times more than in 2004 (only 3054 cases). In addition, when HP filter is used to remove the short-term monthly impact of AIDS incidence to analyze long-term trend and periodicity, a trough is usually observed in January and February of each year, while the case is frequently reported in December of each year. In the remaining months, the number of reported cases remained stable and the volatility has become more pronounced since 2010 (Figure 7). In addition, in the past 14 years, the number of reported AIDS cases has increased significantly, with an annual increase of about 14.384%.

Figure 7 is a time series image of the monthly AIDS medical records from 2004 to 2017. From the figure, we can see that the incidence of cases is gradually increasing. The rate of increase from 2004 to 2010 is relatively slow, and the rate of increase from 2011 to 2012 is suddenly increasing, but the trend of increase from 2013 to 2017 is generally consistent with that from 2004 to 2010.

From January 2004 to April 2018, there were a total of 458,078 AIDS cases, with an average of 2664 cases per month and a standard error of 143 cases. The average annual incidence rate was 2 per 100,000 people. In 2017, the number of AIDS cases reported each month in mainland China reached 57,194, which was more than 18 times the number in 2004

TABLE 2. Design values of predicted network parameters.

Parameter	Parameter value
Input step number	6
Input step size	16
Number of samples per batch	30
optimization method	RMSprop
Learning rate	0.001
Dropout probability	0.6

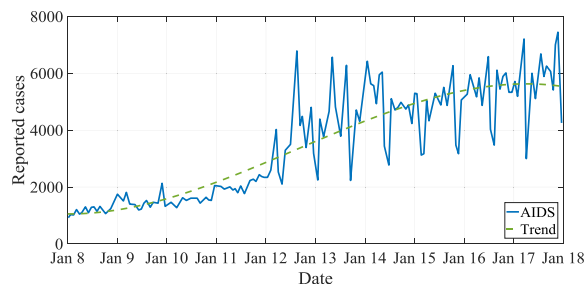


FIGURE 8. Sequence of monthly HIV incidence reported and filtered using HIV filtering method from January 2004 to January 2018.

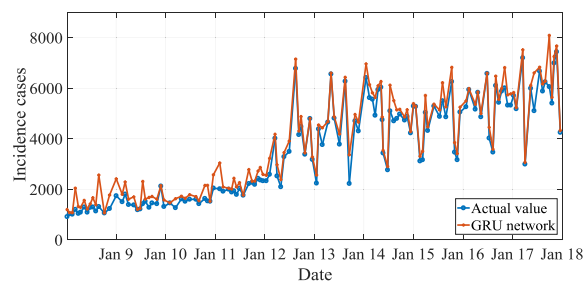


FIGURE 9. Actual and predicted curve of AIDS incidence from January 2009 to January 2018.

(only 3054 cases). In addition, when HP filtering is used to remove the short-term monthly effects of AIDS incidence to analyze long-term trends and periodicity, a trough is usually observed in January and February each year, and the case is frequently reported in December each year. The number of notified cases remained stable for the remaining months, and this fluctuation has become more pronounced since 2010. In addition, the number of reported AIDS cases has shown a significant upward trend over the past 14 years, and has increased by approximately 14.384% per year.

The above figure is a time series image of the incidence of AIDS from January 2004 to January 2018. As shown in the figure, the incidence of AIDS shows a gradual upward trend, which seems to be more obvious since 2010, with a downward trend in 2018. In addition, periodic fluctuations also occur in the same period.

For the prediction experiment of AIDS incidence based on MHP SO-GRU network, the model parameter settings are shown in Table 2.

For the actual change curve and prediction curve of AIDS incidence from January 2017 to April 2018, see Figure 9. MAPE is predicted to be 8.15%. The actual and

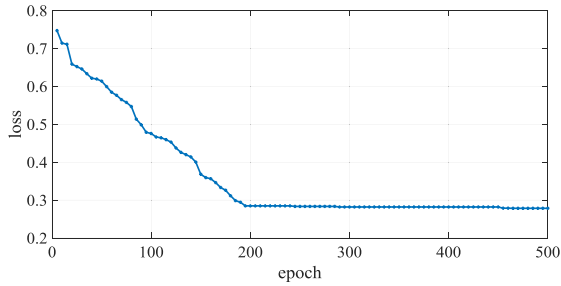


FIGURE 10. Training loss function iteration diagram of AIDS incidence prediction model.

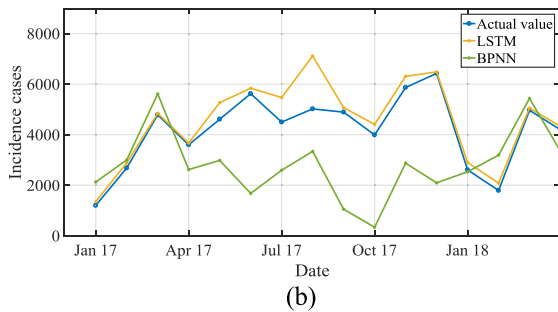
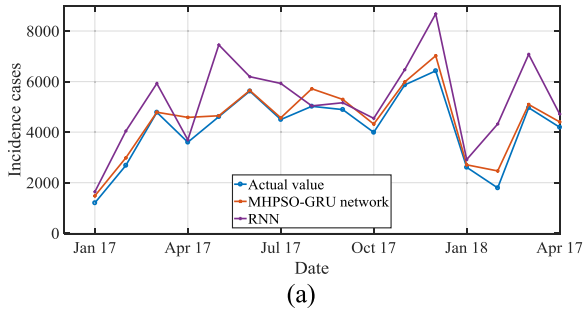


FIGURE 11. Comparison of estimated incidence between four selected models and observations from January 2017 to April 2018.

predicted curves of AIDS incidence from January 2009 to January 2018 are shown in Figure 9, of which the predicted MAPE is 8.21%. From the comparison of the results in Figure 9 and Figure 11, it can be seen that the difference between the whole prediction curve and the actual curve is not big, and the whole prediction error is quite stable within one or ten years, without large deviation, which proves that the method in this paper is effective.

Figure 10 shows the image of the training loss function of the AIDS incidence prediction model changing with the number of iterations. From the above figure, it can be seen that the MHP SO-GRU based AIDS incidence prediction model tends to be stable when the number of training reaches 400, and then the loss function value is close to 0.28, indicating that the prediction effect is better.

C. PERFORMANCE COMPARISON

Table 3 shows the prediction ability and efficiency of the four best fit models selected. We found that compared with MAE, MAPE, mer and RMSE, the fitting results of basic

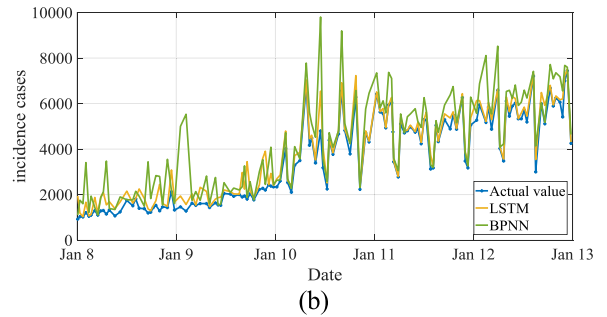
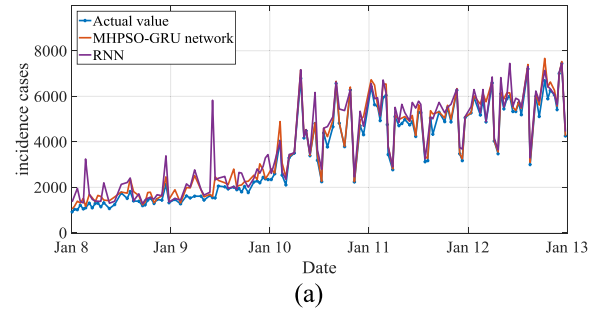


FIGURE 12. Comparison of simulated and estimated incidence cases between selected four models and observations.

BPNN model were the lowest, so as to analyze the future epidemic mode of known cases of AIDS. In the out of sample prediction, AIDS had the maximum value of the above four indicators. The MHP SO-GRU model is considered to be the most suitable model for prediction in the derivative model based on the full consideration of the performance simulated and predicted by these models (Figure 11 and Figure 12). Subsequently, this hybrid approach was re established and data from the whole report were used to simulate the time behavior of AIDS in the near future. As can be seen from table 7, it shows that the disease seems to be on the rise from May 2018 to December 2020.

Figure 11 shows the predicted incidence of AIDS by MHP SO-GRU model, LSTM model, RNN model and BPNN model. It can be seen that BPNN model has the worst prediction effect and the largest deviation from the actual data. The prediction error of RNN model is large, that of LSTM model is small, and that of MHP SO-GRU model is minimum.

Figure 12 shows the global prediction of AIDS incidence by four algorithms: BPNN model, RNN model, LSTM model and MHP SO-GRU model. The blue line is the time sequence of actual incidence of AIDS, the red line is the prediction sequence of MHP SO-GRU model, the yellow line is the prediction sequence image of LSTM model, the purple line is the prediction sequence of RNN model, and the green line is the prediction sequence of BPNN model. In addition to BPNN model, other models have good prediction effect, but MHP SO-GRU model has better prediction effect. As described above, the curve (solid red line) fitted by the MHP SO-GRU method can better approximate the reality (solid blue line) compared to other curves.

TABLE 3. Comparison of fitting performance in samples of the model.

Mod els	Fitting performance				Predicted performance			
	MAE	MAPE	MER	RMSE	MAE	MAPE	MER	RMSE
MHP SO- GRU	50.671	0.037	0.035	148.089	278.875	0.167	0.076	771.710
LST M	54.884	0.053	0.022	142.481	1628.67 9	0.337	0.350	2019.439
RNN	149.388	0.132	0.059	224.292	539.027	0.144	0.116	799.314
BPN N	335.568	0.193	0.126	502.502	1013.00 6	0.290	0.218	1301.985

MHP SO-GRU, MHP SO-Gated Recurrent Unit; LSRM, Long Short-Term Memory; RNN, Recurrent neural network; BPN N, Back Propagation Neural Network; MAE, Mean absolute error; MAPE, mean absolute percentage error; mer, mean error rate; RMSE, root mean square error.9

TABLE 4. Forecasts with the best-performing ARIMA-NAR hybrid techniqnue between May 2018 and December 2020.

Date	Forecasts	95% confidence limits	Date	Forecasts	95% confidence limits
May-18	5057	[2698, 7719]	Sep-19	6591	[2496, 14407]
Jun-18	5943	[3748, 9768]	Oct-19	5355	[2319, 12723]
Jul-18	5418	[3270, 9197]	Nov-19	7311	[2522, 16548]
Aug-18	5521	[3058, 9067]	Dec-19	8092	[2987, 19082]
Sep-18	5401	[2758, 9391]	Jan-20	3445	[680, 8161]
Oct-18	5147	[2955, 8802]	Feb-20	4214	[846, 8730]
Nov-18	6913	[3408, 11317]	Mar-20	7277	[2244, 17095]
Dec-18	7813	[3182, 12322]	Apr-20	6366	[842, 14496]
Jan-19	3115	[1221, 5468]	May-20	6794	[1944, 18073]
Feb-19	2956	[1563, 5963]	Jun-20	8384	[2145, 21143]
Mar-19	6008	[3081, 11567]	Jul-20	7322	[2081, 20386]
Apr-19	4961	[2423, 10065]	Aug-20	6954	[2031, 20488]
May-19	6038	[1966, 11476]	Sep-20	7417	[1795, 21785]
Jun-19	7320	[2890, 14091]	Oct-20	5870	[1801, 19269]
Jul-19	7493	[2802, 13673]	Nov-20	9064	[1759, 25371]
Aug-19	6275	[3018, 13938]	Dec-20	10099	[2398, 29549]

Table 4 shows the prediction value and 95% prediction interval of 2018-2020 based on MHP SO-GRU model, and figure 13 shows the prediction value and prediction interval of 2018-2020. From the comprehensive table and figure, we can see that the incidence of AIDS in the future is still rising, and the situation of AIDS prevention and control is still very serious.

To sum up, according to the longitudinal data of known cases of AIDS reported, we use four useful methods, BPN N, RNN, LSTM, MHP SO-GRU, to model and solve. Figure 12 is a comparison chart of the simulated and estimated incidence rate between the selected four models and observations. The red line is the prediction sequence of the MHP SO-GRU model, and the blue line is the actual incidence rate curve. It is seen that the error of the red line and the blue line is smaller. Table 3 shows the prediction performance comparison of sample fitting of the four models. Mae of MHP SO-GRU network model is 278.875, MAPE is. 0.167, Mer is 0.076, RMSE is 771.710, which is the best of the

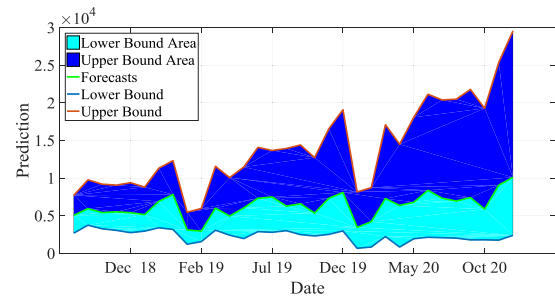


FIGURE 13. Prediction value and prediction interval of MHP SO-GRU model.

four effective methods. This may be a useful technology for estimating the future behavior of AIDS.

VI. CONCLUSION

In the past decades, the incidence rate and mortality rate of AIDS has been increasing. The situation of countries and regions in the world has worsened. The premise of effective containment and intervention strategies for AIDS by relevant organizations is accurate identification and early warning of the future long-term track of this disease behavior.

The advantages of our research include the amazing findings gathered from the AIDS incidence rate data collected from January 2004 to April 2018. The validation of these data has been supported by China’s mandatory infectious disease surveillance system. It can be seen that figure 8 shows that the incidence rate of AIDS in China is increasing. This work can study the future trend of AIDS incidence rate in China, and provide corresponding treatment strategies and intervention measures. In the future work, we plan to study the trend of AIDS rate in other areas to prove that our model is effective and universal.

Even so, there are some limitations in this work. First, the data cover the period including policy interventions (2010). Although the results of the first choice model are satisfactory, the actual impact of the model’s accuracy before and after the sudden rise in incidence rate of AIDS is still unknown. Secondly, in the construction of RNN model, due to the lack of mature theoretical guidance on the optimal number of hidden cells, feedback delay and other key parameters. Although mhps o algorithm is used to improve the number of singular numbers in this paper, its specific simulation and prediction process is largely unknown. Third, it is estimated that 32.4% of AIDS patients in China do not know their infection status. Therefore, the actual situation may be more serious than the estimated situation. Fourthly, many complex factors, including unpredictable factors, may lead to the occurrence and spread of AIDS. The prediction model we built only relies on the known AIDS cases in the past, without considering other factors to predict the unknown cases. Therefore, if conditions permit, factors related to the occurrence and development of AIDS are expected to be used together to assess the epidemic. Finally, the prediction model is based on the cases in China from January 2004 to April 2018. Therefore, the analysis of

the epidemic only represents the situation and trend of AIDS in the country. After being recorded, this method has great benefits for other regions and infectious diseases.

To sum up, we use BPNN, RNN, LSTM respectively according to the longitudinal data of known AIDS cases reported, Four useful methods of MHP SO-GRU are used to model and solve the problem. It is found that the performance of MHP SO-GRU network model is the best. This may be a useful technique to estimate the temporary behavior of AIDS, so it can help health policy makers to reasonably allocate health resources The prevention and control plan of the disease shall be properly formulated. In addition, the number of new HIV infections will continue to increase and is already high, and these issues deserve urgent strategic attention in the context of properly adopted policies.

REFERENCES

- [1] C. M. Brauner, "Heterogeneous viral environment in a HIV spatial model," *Discrete Continuous Dyn. Syst. B*, vol. 15, no. 3, pp. 545–572, May 2017.
- [2] Z. Mukandavire, "Mathematical analysis of a model for HIV-malaria co-infection," *Math. Biosci. Eng. Mbe.*, vol. 6, no. 2, pp. 333–362, Apr. 2009.
- [3] B. M. Adams, "Dynamic multidrug therapies for HIV: Optimal and STI control approaches," *Math. Biosciences Eng. Mbe.*, vol. 1, no. 2, pp. 223–241, Sep. 2004.
- [4] Y. Shen and R. Li, "What do we know from clinical trials on exercise and Alzheimer's disease?" *J. Sport Health Sci.*, vol. 5, no. 4, pp. 397–399, Dec. 2016.
- [5] GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015," *Lancet*, vol. 388, no. 10053, pp. 1545–1602, Oct. 2016.
- [6] R. Siliciano, "A-101 Defective proviruses rapidly accumulate during acute HIV-1 infection," *J. Acquired Immune Deficiency Syndromes.*, vol. 74, no. 34, p. 34, Aug. 2017.
- [7] E. Mukasa, H. Kimaro, A. Kiwanuka, and F. Igira, "Challenges and strategies for standardizing information systems for integrated TB/HIV services in tanzania: A case study of kinondoni municipality," *Electron. J. Inf. Syst. Developing Countries*, vol. 79, no. 1, pp. 1–11, Mar. 2017.
- [8] J.-H. Wu, W. Wei, L. Zhang, J. Wang, R. Damasevicius, J. Li, H.-D. Wang, G.-L. Wang, X. Zhang, J.-X. Yuan, and M. Wozniak, "Risk assessment of hypertension in steel workers based on LVQ and Fisher-SVM deep excavation," *IEEE Access*, vol. 7, pp. 23109–23119, 2019.
- [9] W. Gao, "Sexual Practices and the Prevalence of HIV and Syphilis among Men Who Have Sex with Men in Lanzhou, China," *Jpn. J. Infectious Diseases.*, vol. 68, no. 5, pp. 370–375, Sep. 2015.
- [10] M. Li, Y. Shen, X. Jiang, Q. Li, X. Zhou, and H. Lu, "Clinical epidemiology of HIV/AIDS in China from 2004–2011," *BioScience Trends*, vol. 8, no. 1, pp. 52–58, Mar. 2014.
- [11] S. M. Kariuki, P. Selhorst, K. K. Ariën, and J. R. Dorfman, "The HIV-1 transmission bottleneck," *Retrovirology*, vol. 14, no. 1, p. 22, Dec. 2017.
- [12] F. Zhang, H. Zhu, Y. Wu, Z. Dou, Y. Zhang, N. Kleinman, M. Bulterys, Z. Wu, Y. Ma, D. Zhao, X. Liu, H. Fang, J. Liu, W.-P. Cai, and H. Shang, "HIV, hepatitis B virus, and hepatitis C virus co-infection in patients in the China National Free Antiretroviral Treatment Program, 2010–12: A retrospective observational cohort study," *Lancet Infectious Diseases.*, vol. 14, no. 11, pp. 1065–1072, Nov. 2014.
- [13] J. Jia and G. Qin, "Stability analysis of HIV/AIDS epidemic model with nonlinear incidence and treatment," *Adv. Difference Equations.*, vol. 2017, no. 136, pp. 1–3, May 2017.
- [14] W. Yawen, S. Zhongzhou, Y. Baohu, and Y. Yin, "The application of ARIMA model and ARIMA-GRNN model in the prediction of AIDS incidence," *Chin. J. Disease Control.*, vol. 22, no. 12, pp. 91–94, Dec. 2018.
- [15] K. Yang and P. Qiu, "Spatiotemporal incidence rate data analysis by nonparametric regression," *Statist. Med.*, vol. 37, no. 13, pp. 2094–2107, Jun. 2018.
- [16] M. Masiá, "Prediction of long-term outcomes of HIV-infected patients developing non-AIDS events using a multistate approach," *PLoS ONE.*, vol. 12, no. 9, pp. e0184329, Sep. 2017.
- [17] S. Rezaei, "Global prevalence of depression in HIV/AIDS: A systematic review and meta-analysis," *Supportive Palliative Care.*, vol. 9, no. 4, pp. 404–412, Dec. 2019.
- [18] P. I. Adamu, P. E. Oguntunde, H. I. Okagbue, and O. O. Agboola, "On the epidemiology and statistical analysis of HIV/AIDS patients in the insurgency affected states of Nigeria," *Open Access Macedonian J. Med. Sci.*, vol. 6, no. 7, pp. 1315–1321, Jul. 2018.
- [19] A. Deribew, "The burden of HIV/AIDS in ethiopia from 1990 to 2016: Evidence from the global burden of diseases 2016 study," *Ethiopian J. Health Sciences.*, vol. 29, no. 1, pp. 859–868, Jan. 2019.
- [20] A. Yang, Y. Li, C. Liu, J. Li, Y. Zhang, and J. Wang, "Research on logistics supply chain of iron and steel enterprises based on block chain technology," *Future Gener. Comput. Syst.*, vol. 101, pp. 635–645, Dec. 2019.
- [21] D. Zang, "Long-term traffic speed prediction based on multiscale Spatio-Temporal feature learning network," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3700–3709, Oct. 2019.
- [22] H. Fu, G. Manogaran, K. Wu, M. Cao, S. Jiang, and A. Yang, "Intelligent decision-making of Online shopping behavior based on Internet of Things," *Int. J. Inf. Manage.*, vol. 50, pp. 515–525, Feb. 2020.
- [23] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.
- [24] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000.
- [25] A. Yang, Y. Zhuansun, Y. Shi, H. Liu, Y. Chen, and R. Li, "IoT system for pellet proportioning based on BAS intelligent recommendation model," *IEEE Trans Ind. Informat.*, to be published.
- [26] Y. Ren, C. Teng, F. Li, B. Chen, and D. Ji, "Relation classification via sequence features and bi-directional LSTMs," *Wuhan Univ. J. Natural Sci.*, vol. 22, no. 6, pp. 489–497, Dec. 2017.
- [27] A. Yang, C. Zhang, Y. Chen, Y. Zhuansun, and H. Liu, "Security and privacy of smart home systems based on the Internet of Things and stereo matching algorithms," *IEEE Internet Things J.*, to be published.
- [28] A. F. Agarap, "A neural network architecture combining Gated Recurrent Unit (GRU) and Support Vector Machine (SVM) for intrusion detection in network traffic data," *Assoc. Comput. Mach.*, vol. 10, no. 5, pp. 26–30, Feb. 2018.
- [29] H.-G. Kim and J. Y. Kim, "Acoustic event detection in multichannel audio using gated recurrent neural networks with high-resolution spectral features," *ETRI J.*, vol. 39, no. 6, pp. 832–840, Dec. 2017.



XIAOMING LI received the B.Sc. degree from North China Coal Medical University, in 2006, and the M.Sc. degree from Hebei United University, in 2012. He is currently pursuing the Ph.D. degree with the North China University of Science and Technology. He is also a Lecturer with the North China University of Science and Technology. His main research interests include data mining, occupational epidemiology, occupational health and occupational medicine, and so on.



XIANGHUI XU received the B.Sc. degree from North China Coal Medical University, in 2004, and the M.Sc. degree from Hebei United University, in 2014. She is currently an Attending Physician of the Hospital of North China University of Science and Technology. Her main research interests include prevention and control of infectious diseases, epidemiology of infectious diseases, health education and health promotion, and so on.



JIE WANG was born in Yuncheng, Shanxi, China, in 1995. She graduated from Hainan Medical University. She is currently a Graduate Student with the School of Public Health, North China University of Science and Technology, studying in public health and preventive medicine, mainly researches occupational epidemiology.



SHENG QIN was born in Langfang, Hebei, China, in 1995. He graduated from the Chengdu University of TCM. He is currently a Graduate Student with the School of Public Health, North China University of Science and Technology, studying in public health and preventive medicine, mainly researches occupational epidemiology.



JING LI was born in Handan, Hebei, China, in 1993. She graduated from the North China University of Science and Technology. She is currently a Graduate Student with the School of Public Health, North China University of Science and Technology, studying in public health and preventive medicine, mainly researches occupational epidemiology.



JUXIANG YUAN graduated from the Tangshan Coal Mine Medical College, in 1978. He received the Diploma degree in occupational health and safety from the McMaster University Medical Center, Canada, in 1985, and the Ph.D. degree from the Peking Union Medical College, in 2000. He is currently a Professor with the North China University of Science and Technology. His main research interest is occupational epidemiology.

...