Journal of Autonomous Intelligence (2024) Volume 7 Issue 4
doi: 10.32629/jai.v7i4.651

ORIGINAL RESEARCH ARTICLE

Blockchain-based early warning system for infectious diseases: Integrating risk metrics for complex networks

Yimin Cai[1,2], Jie Ma[1,2], Jasni Mohamad Zain[1,2,*], Yisong Wang[3], Maoxing Zheng[3]

[1] *Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*

[2] *School of Computing Sciences, College of Computing, Informatics and Media, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*

[3] *School of Computer Sciences, Baoji University of Arts and Sciences, Baoji 721000, China*

*** Corresponding author:** Jasni Mohamad Zain, jasni67@uitm.edu.my

ABSTRACT

The emergence of infectious diseases poses a significant threat to public health and requires prompt action to minimize their impact. Traditional infectious disease early warning systems (EDWS) face numerous challenges, including delays in reporting, data privacy concerns, and lack of transparency. Blockchain technology provides a promising solution to these challenges by offering a decentralized, secure, and transparent platform for data sharing and analysis. This paper analyzes cases, medical resources, pathological characteristics, and their connections using hospital data. It proposes an infectious disease risk measurement algorithm based on complex network modeling and integrates the risk measurement algorithm to construct a blockchain-based infectious disease risk early warning and data sharing system. The solution enables the safe storing and sharing of data using blockchain technology; performs distributed global model training via federated learning; and enables rapid and accurate infectious disease risk early warning by integrating smart contracts with physician expertise. Automatic reporting enhances infectious disease risk early warning and reporting.

Keywords: infectious diseases; EDWS; blockchain technology

ARTICLE INFO

Received: 23 May 2023
Accepted: 6 June 2023
Available online: 19 February 2024

1. Introduction

The emergence of infectious diseases poses a significant threat to public health, as highlighted by the recent outbreak of COVID-19. Early detection and response are critical to minimizing the impact of infectious diseases, but traditional early warning systems face numerous challenges, including delays in reporting, data privacy concerns, and lack of transparency[1]. Identifying, notifying, and acting on public health issues as quickly as possible is now a top priority and problem for preventing and controlling illness.

In order to achieve early detection, early intervention, and early control of severe infectious illnesses, China created a massive infectious disease early warning and control system from scratch during the 2003 SARS outbreak. However, this disease control strategy did not play the predicted role in the 2020 new crown pneumonia outbreak. In the infectious disease early warning and control system in place in China, infectious illness prevention and control is a process of "diagnosis-reporting-summary". Before a suspected infectious condition can be reported, it must be diagnosed. On the one hand, it takes a certain amount of time and requires the doctor's knowledge to detect new infectious illnesses， SARS prevention and treatment[2]. On the other hand, physicians and hospitals are under increased pressure to report specific infections when they are suspect, delaying the optimal period for early warning. Wuhan City missed at least four early warning chances for the new Chinese pandemic[3].

The old approach of reporting based on the outcomes of individual case diagnoses is unreliable for the early warning of emerging serious infectious illnesses. It is required to build a collaborative early warning and reporting system for infectious illnesses, transitioning from a summary of individual cases to a summary based on the characteristics of group cases and improving the data. reporting routes and platforms[4] for the formation of an effective emergency data management system. Likewise, the collaborative early warning and reporting of infectious illnesses is based on a good hospital data sharing method for epidemic risk. Medical data is sensitive to privacy, and the sharing procedure should pay particular attention to data security and authority control. As a distributed accounting system, blockchain features decentralization, transparency, openness, anonymity, and immutability. The infectious disease risk early warning system based on blockchain technology can provide secure, trustworthy, and tamper-proof data storage. Protect people’s privacy and sensitive information in medical data; set up smart contracts and consensus algorithms to automate early warning and reporting; and make it easier to track the whole early warning process.

To address this gap, this paper proposes a novel Blockchain-Based Early Warning System for infectious diseases that integrates risk metrics for complex networks. The proposed system leverages the power of blockchain technology to provide a secure and transparent platform for data sharing and analysis. Moreover, it utilizes risk metrics to quantify the probability and severity of infectious disease outbreaks, enabling timely and effective responses. This paper designs a complex network-based infectious disease risk early warning algorithm based on hospital operation data and case information, integrates this early warning algorithm to build an infectious disease risk early warning system based on blockchain and complex networks, and trains the global model through federated learning. Using smart contracts and consensus mechanisms to achieve early warning and automatic reporting of infectious disease risks, breaking the original symptom diagnosis mechanism of individual cases, turning to the monitoring and discovery of group abnormalities, supporting joint judgment and identification of multiple hospitals, and based on blockchain technology, protect privacy, store and share data securely, and come up with new ideas and solutions for setting up early warning systems and algorithms for infectious diseases.

2. Research status of infectious disease early warning methods and systems

2.1. Basic principles and main models of infectious disease early warning

(1) Basic principles:

Early warning of infectious illnesses refers to the issuing of warning signals before the outbreak or epidemic of infectious diseases, or at an early stage, to remind the outbreak or epidemic that the outbreak or epidemic may occur or that its scope may extend[5,6]. There are many international terms for infectious illness early warning. The most often used are: epidemic detection, aberrancy detection, and early warning (early warning).

In a certain period, place, and population, the frequency of infectious illnesses varies within a defined range; that is, it is kept at a level consistent with normality. When the incidence of infectious illnesses reaches this threshold, an abnormal circumstance has developed. This further raises the probability of an infectious illness breakout or epidemic. The fundamental premise of infectious disease early warning is to locate and identify abnormal circumstances of infectious illnesses that exceed the predicted normal level based on surveillance data pertaining to infectious diseases using certain analytical techniques or models[7]. Infectious disease early warning is a key part of public health emergency response. It is based on infectious disease surveillance and has a high level of sensitivity and timeliness[6].

(2) Main early warning model:

The early warning process of infectious illnesses may be seen as a process of information transformation; that is, monitoring data is converted into early warning information, of which the early warning model is a crucial component[8]. Infectious illness early warning models are primarily concerned with the detection and interpretation of monitoring data from the temporal dimension. With the gradual development of spatial statistical methods, a method for detecting and warning of the abnormal aggregation and distribution of infectious diseases in the spatial dimension has been established, and spatiotemporal early warning models of infectious disease outbreaks or epidemics based on different data sources have been explored.

The time-early warning model focuses on the temporal distribution or variation features of infectious disease-related surveillance indicators in a particular region so as to reflect whether the danger of infectious disease outbreaks or epidemics has grown considerably. The statistical process control approach based on quality management has been used extensively in public health monitoring[9–16], and the moving percentile method, cumulative sum (cumulative sum) method has been developed a moving level control chart technique with exponential weighting.

There are many early warning models for infectious diseases, such as the exponentially weighted moving average control chart technique[17–19].

If the current number of instances hits or exceeds the threshold, a warning is generated. Moreover, popular time series analysis approaches, such as the auto-regressive integrated moving average model (auto-regressive integrated moving average model), examine many aspects of time series changes (trend, periodicity, and seasonality, etc.) and anticipate a certain period. Risk analysis and early warning of infectious illnesses[20] have also investigated and used the epidemic level of infectious diseases throughout time. How to supply monitoring data as fast and as accurate as possible has a direct bearing on the application impact of the temporal early warning model. Also, the temporal model can't give accurate early warnings of clustered epidemics in a small area because it doesn't have data about the area.

The spatial early warning model analyzes primarily the spatial distribution or change characteristics of infectious disease-related cases or events at a specific time point or time period and compares the incidence level of the area of interest to the incidence level of all or surrounding areas to detect the incidence of the disease in the area of interest. On the basis of whether there is a statistically significant geographical cluster of the epidemic scenario, an early warning signal should be sent. The foundation of the spatial early warning model is that infectious disease-related data includes geographic information, such as the home address of the reported case or its latitude and longitude. Common techniques of aggregation analysis for spatial early warning models include space scan statistics, Getis-Ord Gi*, and Moran's I and so on[21–23]. The spatial scan statistics provided by Kulldorff[24] and the model computing program SaTScan, created by Kulldorff, have garnered the most interest and applications. Since it has not been compared to the historical incidence level, the spatial model's early warning signal may merely represent the seasonal shift of the epidemic condition and how geographically comparable the cases are. There seems to be no unusual alteration in the time dimension.

The spatiotemporal early warning model considers the temporal and spatial variations of infectious disease outbreaks. By mining and using the time and location information included within infectious disease surveillance data, it is possible to detect high-risk regions and times of infectious disease outbreaks or epidemics, hence enhancing aggregation, the precision and timeliness of sexual detection and early warning. For instance, in the spatiotemporal early warning model of the China infectious diseases automated-alert and response system (CIDARS)[25], different thresholds were first set according to the epidemiological characteristics and historical incidence levels of 17 infectious diseases; the time model (moving percentile method) is used to detect whether there are abnormal changes in the current number of cases in the county (district). For the search unit, the statistically skewed distribution of the current number of cases is utilized.

The space-time scan statistic based on distinct geographical detection windows has been extensively used for spatiotemporal aggregate analysis and early warning of case report data and symptom monitoring-related indicators (school absence, absenteeism)[26–28]. In addition, models of early warning based on elements associated with the spread of infectious illnesses (such as climate, transmission medium, animal populations, human migrations, and economic situations, etc.) have been explored and implemented throughout time. Using the relevant risk factors and data of infectious disease epidemics, for example, a Bayesian-based spatial or spatiotemporal regression model can be built. This model can be used to estimate the epidemic season and high-risk areas of animal-borne and vector-borne infectious diseases, as well as ways to warn of transmission risks (possible changes in the space-time dimension)[29,30].

2.2. Research status of early warning methods for infectious diseases

Infectious disease early warning refers to the dissemination of warning signals prior to and during the early stages of an infectious disease outbreak or epidemic in order to convey the danger that the outbreak or epidemic may occur or that its scope may increase. In order to conduct early detection, early warning, and quick reaction to disease outbreaks, the fundamental premise is to uncover and identify aberrant circumstances of infectious illnesses above the anticipated normal level using infectious disease surveillance data and early warning techniques[31]. Early warning approaches are mostly derived from time and spatial dimensions, using qualitative methods such as cumulative summation (CUSUM), ratio map method, Delphi method, and time series analysis[32], and autoregressive synthesis[33]. Quantitative techniques such as the moving average (ARIMA) model[33] and the Holt-Winter algorithm[34] are used to assess early warning, but the accuracy and timeliness of early warning are inadequate. As a result of technological advancements, diverse infectious disease monitoring systems have emerged and the sources of infectious disease early warning data have varied. In addition to conventional clinical data, it also incorporates multi-source data such as search engine data, social media data, and e-commerce sales data. Google successfully predicted influenza outbreaks in 2009 using search phrase data, and their prediction was more than a week earlier than the data from the Centers for Disease Control and Prevention[36]. With the successful application of machine learning in various fields, machine learning-based prediction models have achieved success in predicting infectious diseases. Jia et al.[37] established a linear model, a time series analysis model, a boosting tree model, and a deep learning model (RNN) based on historical morbidity, mortality, search engine, and seasonal data. The experimental findings demonstrate that the RNN model performs better. Using an artificial neural network technique using environmental, meteorological, and epidemiological data from the National Health Research Institute (SIVIGILA) and the Abra Valley Early Warning System (SIATA)[37], developed an early warning system for dengue disease and empirical test. In conclusion, in the early stages, the early warning methods of infectious diseases are primarily focused on time series analysis, the autoregressive comprehensive moving average model, and other methods. However, in the near future, they will primarily focus on the role of machine learning, neural networks, and other technologies in the early warning of infectious diseases as well as the history of infectious diseases. Concerning data and transmission dynamics, there is a dearth of studies on the network-based connections between verified cases.

2.3. The development status and deficiencies of China’s infectious disease early warning system

In order to apply the early warning technique of infectious illnesses to the monitoring of infectious disease outbreaks or epidemics, it is necessary to build an early warning system that can be implemented immediately on the basis of early warning technology. Early in the 1990s, nations such as the United States started exploring the implementation of monitoring systems based on infectious illness symptoms. China’s infectious disease symptom monitoring and early warning system has expanded fast since the 2003 SARS crisis. China’s network direct reporting system for infectious diseases and public health emergencies (referred to as the network direct reporting system) was launched nationwide in April 2004. The national infectious disease monitoring automatic early warning (space-time model) system was used as a pilot project in 2008. The work has officially begun[38,39]. To a certain extent, these systems have solved the problems of information reporting of infectious diseases and public health emergencies as well as early monitoring and early warning of infectious diseases, and they have played a significant role in the early warning and prevention of infectious diseases in my country. Even if the existing infectious disease early warning system in China is largely comprehensive, there are still limitations in early warning speed, early warning methodologies, and information exchange, particularly in the face of new and significant infectious illnesses. The infectious disease early warning and surveillance system has the following issues:

First, the early warning time is inadequate. China's present strategy for early detection of infectious illnesses consists mostly of diagnosis followed by reporting, primarily for communicable infectious diseases. Although "suspected cases" are a possibility, the earliest signs of emerging infectious illnesses may not be obvious, causing warning time to be overlooked. In addition, the present early warning and monitoring systems for infectious illnesses in China are mostly based on regions, are extremely comprehensive, have significant data collection time costs, and have limited capacity to react to developing infectious diseases that are difficult to identify.

Second, the technique of early warning is rather retrograde. Regarding early warning data sources, the present early warning system includes only clinical diagnostic information and historical data on communicable infectious illnesses, and no mechanism for the exchange of information across medical institutions at all levels has been developed. Current early warning models are mostly based on the time model technique and the spatiotemporal model method, which are part of the rule model and can only offer early warning for infectious illnesses that are already recognized. Rapidly evolving technologies such as machine learning, artificial intelligence, and big data analysis have not been used for early warning methods in recent years.

Third, there is an inadequate exchange of information. Multiple institutions, organizations, and departments must coordinate their activities in order to detect and respond promptly to infectious illnesses. However, in the current early warning system, data exchange between medical institutions and between medical institutions and the CDC, government, and other agencies is inadequate. There isn't much sharing of real-time information about the risk of infectious diseases, which makes it harder for hospitals and other institutions to make decisions and give early warnings about the spread of infectious diseases.

2.4. Application of complex network in infectious disease early warning research

A complex network is a large-scale network with a complicated topology that abstracts people or complex systems in the objective world as nodes and the interactions between nodes as edges. These nodes and the edges that link them constitute a complicated network for a structured representation[40–43]. Unlike conventional static approaches, sophisticated network modeling may dynamically capture network abnormalities. The transmission of infectious illnesses through complex networks has been a popular area of study. To simulate the dynamic process of infectious diseases in a network, researchers have studied a number of infectious disease propagation models, such as the SIS model[44], SIR model[45], and SIRS model[46]. The study of these transmission models plays a crucial role in controlling and predicting the spread of infectious diseases. Li et al.[47] compiled and validated a hand, foot, and mouth disease epidemic prediction and early warning model using the SIS model and a curve fitting approach based on complex network theory. Other researchers use the structural properties of complex networks to anticipate the emergence of large-scale population illnesses by monitoring, identifying, and evaluating a subset of network members with a high infection risk. Chrisctakis et al.[49] suggested an early detection technique for static contact networks in 2010, picking random people's acquaintances as monitoring objects and accurately predicting the influenza epidemic at Harvard University.

However, this method's assumption is based on the static contact network concept. However, in the current world, the contact network topology between individuals exhibits dynamic change[49]. Holme et al.[50] analyzed the challenge of early identification of infectious illnesses within a dynamic interaction network topology. Using the dynamic immunological approach as a model, he made a method for early diagnosis of infectious diseases based on a network of interactions that changes over time and then tested it in the real world.

The standard spatiotemporal model is better at spotting infectious diseases early than the complex network model, but more research needs to be done on how well it can spot new infectious diseases early.

3. Key technologies for multi-party decision-making fusion monitoring and early warning of infectious diseases based on blockchain

This section proposes a blockchain-based multi-party decision-making fusion monitoring and monitoring of infectious diseases in order to fully concentrate the monitoring forces of the government, medical institutions, and the general public and to implement long-term, efficient, real-time, and stable infectious disease early warning and social information disclosure. Under the premise of security and privacy, this technology can leverage the decentralized data and component sharing market and the epidemic information collaborative monitoring market to facilitate the sharing of multi-source early warning data, the integration of various early warning technologies, and the long-term participation and collaborative mining of multiple monitoring parties. Smart contracts with predefined decision fusion algorithms and aberrant detection criteria may be utilized as continuous real-time monitoring systems to enhance the timeliness of early alerts. The infrastructure, the early warning system, and how the most important technologies will be set up will be shown.

3.1. Infrastructure

Figure 1 depicts the fundamental architecture of the early warning system described in this study. On the blockchain as a whole, blockchain-based multi-party decision-making fusion monitoring and early warning solutions for infectious illnesses are being constructed. User identity, data characteristics, monitoring modes, and the medical institution side of closed source data relies on federated learning alliance detection, while the social side of open source data relies on three sub-modules of collaborative monitoring based on the open blockchain market. This module, which combines smart contracts, allows users of each sub-module to personalize the client. All private files will be encrypted and saved in IPFS during interaction, while only the hash value and all public files will be stored on the blockchain for verification and tracking of the interaction record.

There are five primary smart contract components, including smart contract authority management, smart contract monitoring (open source data), smart contract monitoring (closed source data), smart contract decision-making fusion, and smart contract credit incentive. The group consists of sub-contracts that inherit from and call each other on the blockchain, and define varying use rights and data visibility based on functional needs. The oracle will be a trusted external data source for smart contracts in the system. It will connect to verified off-chain authoritative sites, IPFS addresses, or other data sources of designated blockchains for contracts to ask about the state of the outside world, check trigger conditions, and run contracts.

The system includes two early warning channels: one is an epidemic early warning channel issued by administrative agencies after administrative review and approval; the other one is an intelligent contract decision-making fusion component that integrates medical, government, and social decision-making information. Institutional and social decision-making information will be evaluated, and an epidemic early warning channel will be automatically activated if the early warning requirements are satisfied. In order to avoid unknown illnesses, the decision rationale, early warning level, and health protection advice corresponding to the early warning disease fuzzy search will be revealed concurrently with the release of the early warning channel. have a role in early prevention.

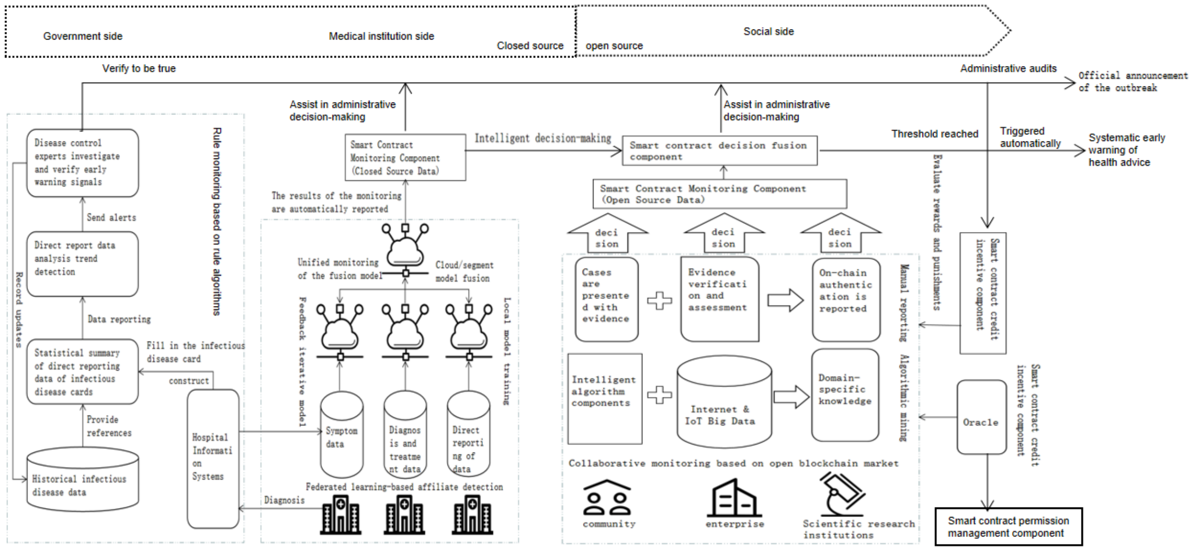

Figure 1. Key technology framework for multi-party decision-making fusion monitoring and early warning of infectious diseases based on blockchain.

3.2. Early warning process

This section will describe the whole operating mechanism and early warning procedure of the early warning technology described in the current study. The particular steps are outlined below. Each port's users conduct epidemic surveillance activities. The government side provides a closed-loop monitoring data interface based on rule algorithms, which consists primarily of the following: standardized infectious disease early warning technology access interface for easy access and replacement of intelligent early warning technology components on the social side; government-side early warning signals. In a unified format, the response and processing information summary interface summarizes the early warning signals monitored by the government and their verification results in real time. The summary interface is linked to the smart contract credit incentive as a way to measure how well medical institutions and social machines make decisions.

The medical institution monitors the federated learning-based alliance. On the basis of private computing architectures such as federated learning, medical institutions train collectively to obtain a unified shared monitoring model without exchanging their own private data. This model is used to monitor the epidemic situation of infectious diseases in multiple medical institutions. It is immediately sent to the smart contract monitoring component (closed source data) through the built-in program. This means that the hospital doesn't have to check for and report infectious diseases by hand.

Based on the open blockchain market, the social side undertakes joint monitoring. The social side constructs a decentralized data and component sharing market and a collaborative monitoring market for epidemic information on the blockchain, and gathers multisource monitoring information via crowdsourcing and cooperation. Each system user has the option of submitting, developing, or validating early warning data sets and intelligent early warning technology components on the decentralized data and component sharing market. Algorithm components and datasets will be encrypted on-demand and deposited into IPFS for index addressing and global sharing, and all data and algorithms submitted to the market will be assigned an algorithm ID and a data ID containing their own details, their quality, and the date they were submitted. Scores are established at the time of first submission and revised by buyer assessment or credit incentive components during future monitoring. The government and medical institutions may acquire algorithm components on demand from the market, thus lowering the complexity of data analysis. The market for collaborative monitoring of epidemic information comprises two monitoring modes: manual monitoring and algorithm monitoring. Any system user may submit manual monitoring data and certification documents to the market. The certification papers must be kept openly in IPFS and allow smart contracts and forecasts. Verification and assessment of machines; algorithm monitoring requires social accounts to offer computer power, freely mix data and algorithms with varying quality scores, monitor the epidemic condition in real-time, and provide monitoring findings to the market. Manual monitoring results and algorithm monitoring results will be assigned a manual decision ID and an algorithm decision ID, recorded in the blockchain, and reported to the smart contract monitoring component (open source data) for fusion. The fusion results are then reported to the smart contract decision fusion component.

(1) Software choice (open source vs. closed source)

Early warning from the Fusion Trigger System smart contract monitoring component (closed source data) or smart contract monitoring component (open source data) is received within a period of time based on decision-making quality, early warning diseases, and early warning scope. The disease's features and monitoring requirements use a number of early warning systems. The results of combining the choices for monitoring infectious diseases are then sent out as an early warning system, and a fuzzy search algorithm is used to find the health advice that goes with the early warning illness.

(2) The smart contract feedback incentive includes a credit incentive component.

Throughout the entire process of early warning, the smart contract credit incentive component will receive and record the response processing information from the government side; the integrated monitoring results from the medical institution side and social side; and the epidemic information from the oracle link authoritative site. To encourage long-term monitoring of the system, the credit and quality of all accounts, institutions, data, and algorithms are evaluated in reverse, based on the judgment results, after the actual outbreak of the epidemic, and the relevant scores, voting weights, rewards, and punishments for contributors are updated.

3.3. Implementation of key technologies

3.3.1. Blockchain-based federated learning

Federated learning is a new architecture for machine learning developed by Google in 2016 to address the issue of data silos and preserve data privacy. To increase model accuracy, they aggregate training data on the premise of trading private data. The architecture of federated learning is especially suited for unified monitoring on the side of medical institutions due to the fact that medical institution data is often difficult to share, and the monitoring of infectious illnesses in the same scope necessitates the use of standard models. Existing frameworks are often set on a centralized cloud or server, posing risks of data loss and single points of failure. In this study, a blockchain-based federated learning paradigm is proposed as a way to solve the problems listed above. Using symptom monitoring-based early warning technology as an illustration, the process of classifying unknown diagnostic data after learning disease characteristics from historical diagnostic datasets using machine learning algorithms can be transformed into a supervised machine learning classification problem. Denote the local symptom database of the *k*-th medical institution containing n_k symptom data samples as $\mathrm{D_k} = \{()\mathrm{x}_1, \mathrm{x}_2, ..., (\mathrm{X}_{\mathrm{n_k}}, \mathrm{Y}_{\mathrm{n_k}})\}$. The empirical risk of the local classification model to the local symptom database is the local classification model objective function. It can be formalized as equation (1)[51]:

$$F_k(w) = \frac{1}{n_k} \sum_{j_k=1}^{n_k} f_{j_k}\left(w, x_{j_k}, y_{j_k}\right) \tag{1}$$

where w is the model parameter.

Then the global model composed of m medical institutions needs to minimize the overall classification error, that is, the global model objective function is formally expressed as follows:

$$\min_{w} F(w), F(w) := \sum_{k=1}^{m} p_k F_k(w), p_k \geq 0, \sum_{k=1}^{m} p_k = 1 \tag{2}$$

Among them, p_k is the weight of the k-th local classification model in the overall model.

At each iteration, the local classification model reads the global model parameters, updates the local model parameters, calculates the model gradient using local data, and communicates the encrypted gradient to the global model. The global model gathers gradients in accordance with the goal function and changes the global model's parameters. Repeatedly multiple times, until the iterative stopping requirements are satisfied, it is possible to acquire an ideal global model that may be used to monitor each local database in reverse for the diagnosis of patient diseases. According to the features of recognized illnesses, the training data for the infectious disease symptom monitoring model is split into CT image data, electronic medical record data, ECG data, and network direct report data, etc. For unknown disorders, the clustering model, principal component analysis, adversarial network, etc. may be used. Anomaly detection techniques are based on semi-supervised or unsupervised learning.

Encrypted transmission and fusion of model parameters are essential for federated learning interactions. To prevent denial or data loss, blockchain-based federated learning may add the interaction process to the blockchain. Depending on the data amount and model quality of each medical institution, each medical institution may also be assessed. In accordance with the contribution value, the fusion weight of the local model in the global model is updated to increase the global model's accuracy, or the economic incentives of medical institutions are fed back to attract additional participants to continue contributing. Two blockchain-based federated learning models, off-chain aggregation and contract aggregation, are supported by the early warning technology suggested in this study. The off-chain aggregation utilizes the chain federated learning methodology described in the study of Kim et al.[52]. After locally computing the model, each medical institution randomly chooses a verification node to submit a transaction including model parameters, data volume, and model running time, and the verification node validates the data volume. After the model's execution time has been matched, the packaged transaction is posted to the blockchain. All parameters of other submodels in the new block are downloaded in real time by the medical institution and aggregated locally. The verification node is compensated based on the number of verification jobs it completes. In this mode, the system must construct a consortia-specific blockchain network and employ cross-chain technology often to anchor consortium chain data on the main chain to maintain data security and training efficiency. A method that is more succinct is contract aggregation. Medical institutions use homomorphic encryption technology to encrypt the local model parameters before uploading the encrypted data to the weight aggregation smart contract. The contract returns the result of calculating the fusion weight. The output acquired by homomorphically encrypted data after specific processing is compatible with the output of the original unencrypted data treated in the same manner after decryption. Thus, the weight parameters remain privately encrypted when the contract fuses the weights. Because smart contracts can only do so much computing, the contract aggregation mode has stricter rules for model types and homomorphic encryption.

3.3.2. Open blockchain market and decision fusion algorithm

The early warning mechanism of the social-end intelligent early warning technology utilizes Internet or Internet of Things big data as the model input, models the expected normal level of a specific infectious disease using the powerful data fitting capability of artificial intelligence, and then monitors, identifies, and discovers the abnormal level. In order to produce the early warning effect, the standard mathematical model must conform to the abnormal circumstances of the standard mathematical model. High-quality early-warning data and credible early-warning models are the keys to accurate early-warning for the social side's early-warning judgments. In this article, **Figure 2** depicts the collaborative monitoring mechanism of the open blockchain market. **Figure 2** depicts the main parameters such as user ID, data ID, algorithm ID, manual decision ID, and algorithm ID that are discussed in this work. IPFS stores shared early warning data, shared intelligent early warning technology, and manual decision-making evidence materials in the system. The smart contract authority management component initializes the credit/quality scores of all users, data, algorithms, and choices, which are then recorded and maintained by the smart contract credit incentive component. Methods for updating include human validation and system review. In user verification, manual decision-making can be verified by any user, whereas data and algorithms can only be verified after the user has obtained permission to use them. In system evaluation, the system treats information based on the response of the government, the fusion results of decision-making on the medical institution, and the social side. The correct early warning result from authoritative sites is the correct early warning result, and the relevant submission decisions in the system are based on these results. After all verification and assessment findings are delivered to the credit incentive component of the smart contract, the credit incentive component will retroactively reward and penalize relevant users and change related data scores based on predetermined rules. For instance, for a correct algorithm decision, the decision submission account, algorithm submission account, and data submission account, etc. can obtain economic incentives and credit score improvement, and related algorithms and data can obtain improved quality scores; economic incentives are issued and circulated in the form of blockchain tokens, and the sources can include penalties for incorrect decisions, penalties for malicious users, and regular government incentives. Digital assets like data and algorithms can be priced using quantitative indicators like credit and quality ratings, and submitters can work with the relevant blockchain data retrieval system and demand matching system to build a decentralized market for sharing data and components.

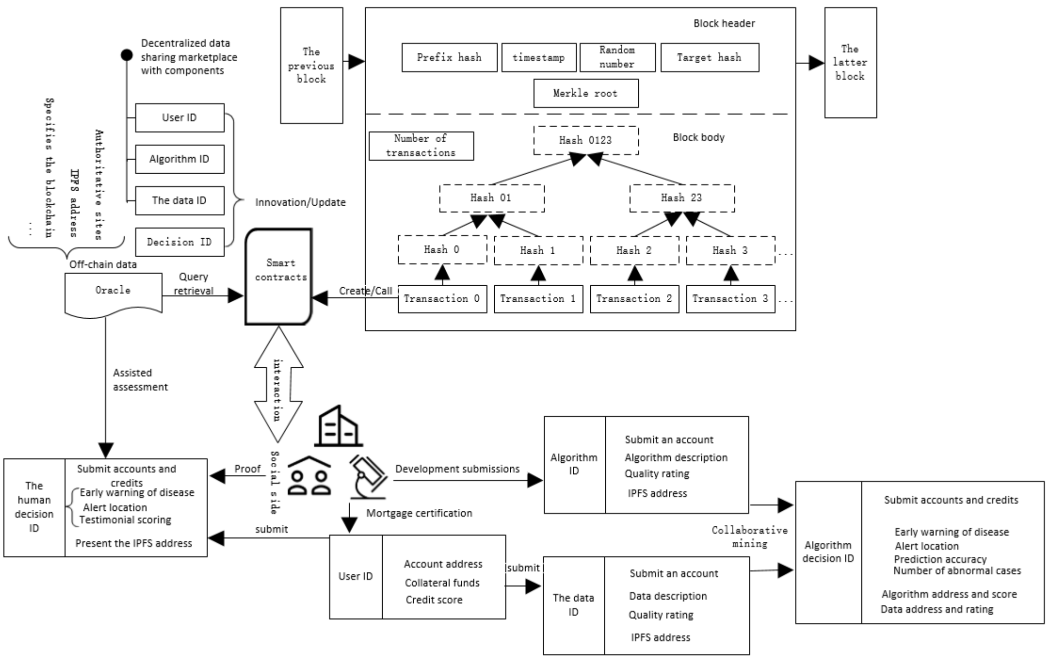

Figure 2. Collaborative monitoring process of open blockchain market.

During processing, the input data of the smart contract monitoring component and the smart contract decision fusion component are structured decision data in a standard format that can be differentiated and fused at the semantic level or weighted at the data level. In particular, the smart contract monitoring component and the smart contract decision fusion component typically consist of two fundamental algorithms: the choice selection algorithm and the decision fusion algorithm, algorithm quality, data quality, and other screening indicators to quickly filter invalid or malicious decisions in massive decisions; decision fusion algorithms set decision fusion rules, and decision fusion rules set manual and algorithmic decisions based on factors such as evidence score, prediction accuracy, and prediction error rate to achieve precise consensus decision-making. On the medical side, it is usually assumed that monitoring results produced from clinical data are more accurate, while monitoring results obtained from social data are timelier. By changing the weight of fusing the two, one can account for how accurate and timely illness monitoring is.

4. Blockchain-based infectious disease risk early warning and data sharing system

As a distributed accounting system, blockchain features decentralization, transparency, openness, anonymity, and immutability. It has several uses in data security, distributed storage, and privacy protection. Using blockchain technology to build an early warning and data-sharing system for infectious disease risk is helpful in the following ways:

Initially, guarantee the security and immutability of data. The data is kept in a distributed way, which effectively ensures data security and integrity and is difficult to lose; the timestamp and encryption mechanism of the blockchain ensure that the data of each block is in chronological order. The hash digest of the information from the previous block is saved in the next block. Information that has already been added to the chain cannot be changed.

Second, safeguard medical information and personal privacy. Medical data comprises personal privacy and is sensitive data. Without divulging actual data, the blockchain may leverage encryption methods and zero-knowledge proof procedures. In the case of early warning for infectious diseases, blockchain technology and federated learning can be used to train the distributed infectious disease risk model without putting any hospital's data at risk of being leaked or made public.

Third, implement distributed consensus and smart reporting. Using smart contracts and a complicated network model of infectious illnesses and hospital data, the infectious disease risk is estimated in real time, and circumstances when the threshold is exceeded trigger automated alerts. Using an acceptable consensus method, the medical institutions in the area agree on the risks and early warnings of infectious diseases and automatically report on this basis. This is called decentralized intelligent joint reporting.

4.1. System architecture

Based on the complex network risk measuring algorithm for infectious illnesses described before, this research develops a blockchain-based early warning and data sharing system for infectious disease risk, as seen in **Figure 3**. In the data acquisition layer, data that may be related to infectious diseases is extracted from the hospital's existing information system and integrated; in the risk measurement layer, the complex network is used to measure the risk of infectious diseases and provide early warning; and the model training layer is based on federated learning to train infectious diseases in the area. The joint model of disease risk is used to determine global parameters and return them to the risk measurement layer; in the data exchange layer, the infectious disease risk data chain is constructed in units of regions, and data exchange and sharing are accomplished using side chain technology; distributed consensus and automated reporting of disease risk.

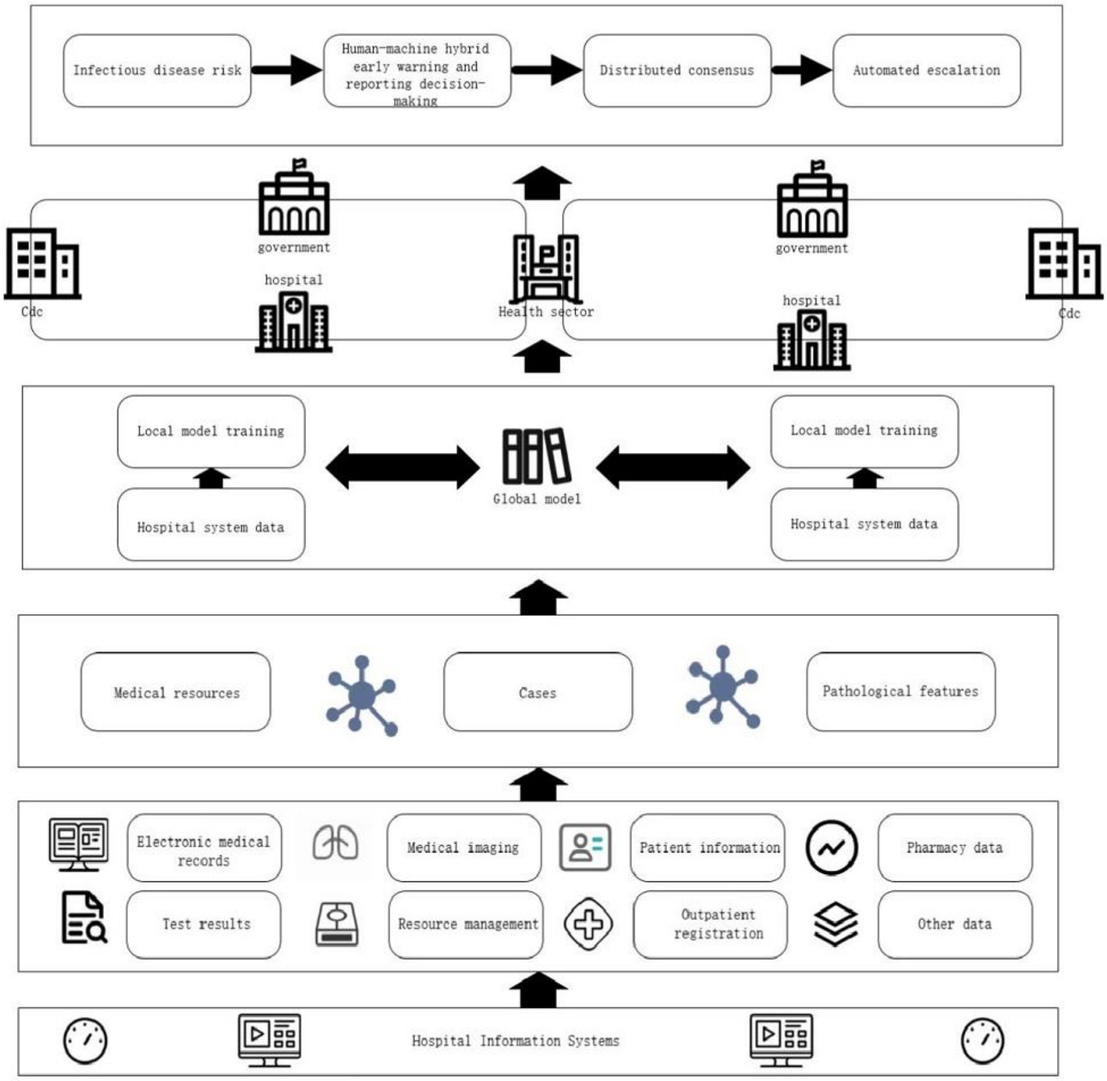

Figure 3. Blockchain-based infectious disease risk early warning and data sharing model.

4.2. Implementation of key technologies

In addition to the infectious disease risk measurement algorithm based on complex network modeling, the system proposed in this paper makes use of the following key technologies: joint model training based on federated learning; multi-level data exchange based on side chain technology; and human-machine hybrid consensus based on smart contracts. Each of these technologies is described in more detail below.

4.2.1. Joint model training based on federated learning

The establishment of thresholds and parameters within the infectious disease risk network model requires ongoing training and modification based on operational and medical record data from each medical facility. However, medical data includes private and sensitive information. In the process of several institutions jointly training models, the protection of data ownership and privacy is crucial. The development of a training model that protects the confidentiality of hospital data is required. Blockchain-based distributed storage and computing, coupled with federated learning, enables sensitive data such as personal privacy in hospitals to be stored and executed as raw data on the local server, while risk models, risk values, medical resource occupancy, and infectious disease reports are shared. For model training, large-scale and comprehensive data, such as cards, are employed, which considerably decreases the danger of data leakage.

Federated learning is a platform for distributed machine learning that ensures data privacy and enables various subjects to accomplish global model training by sharing data models rather than raw data[29]. As seen in **Figure 4**, it is anticipated that hospitals A and B train the infectious disease risk measuring model concurrently. First, each hospital takes case samples relevant to infectious illnesses from the local hospital information system, extracts the characteristics necessary for computation, and then, based on a previous consultation, calculates the incidence of infectious diseases. Align the samples to assist the joint model's development. Each hospital produces the local risk measurement model and uploads it to the blockchain. Each node in the blockchain then exchanges and validates all models, and the smart contract calculates the global model based on a summary of each model. As blocks are generated to hold new local models, parameters will be sent back to each hospital. The hospital then retrains using the new global model parameters and repeats the process until the loss function converges, the final global model is obtained, and the model training process is complete.

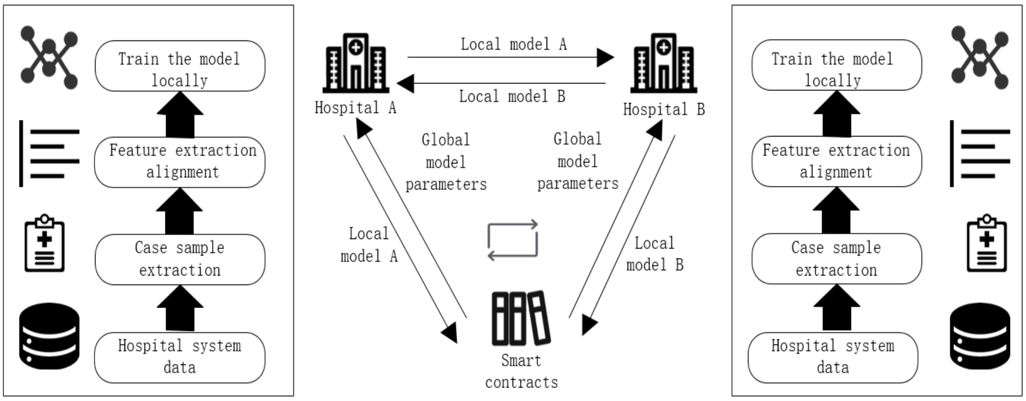

Figure 4. Schematic diagram of federated learning.

4.2.2. Multi-level data exchange based on side chain technology

The building of a countrywide blockchain-based early warning and data-sharing system for infectious disease risk must include a complex institutional framework, including medical institutions of all levels, health administrative agencies, people's governments, and the CDC. Constructing numerous regional-level blockchains segregated by administrative areas helps minimize issues such as data duplication and authority confusion, as compared to building a blockchain to track infectious disease risk across all relevant institutions. However, it is vital to explore how to safely transmit and move data across different blockchains under this arrangement. Sidechain technology facilitates the movement of data across multiple blockchains. It serves as a conduit for connecting many chains for safe interaction. Using side chain technology across different blockchains makes it possible for data to move in a safe and efficient way.

Different chains are created based on administrative areas, with the district and county level serving as the units of data exchange. A consortium chain network consists of all clinics, hospitals, and other medical facilities in a district or county, as well as CDCs and health departments. Each node in this network provides data such as infectious disease risk models, infectious disease risk indexes, and infectious illness report cards based on local hospital data, as well as early warning, decision-making, and consensus data. As depicted in **Figure 5**, the data sharing chains in different regions are connected via side chain technology, and health departments at all levels are used as transit nodes to share district and county-level epidemic data with higher-level departments based on smart contracts in order to achieve a nationwide situation. Data connection for early warning of infectious disease risk. Each institution has access to the data of all of its subordinate institutions and to the combined data of its peers, but not to the data of its parent institution.

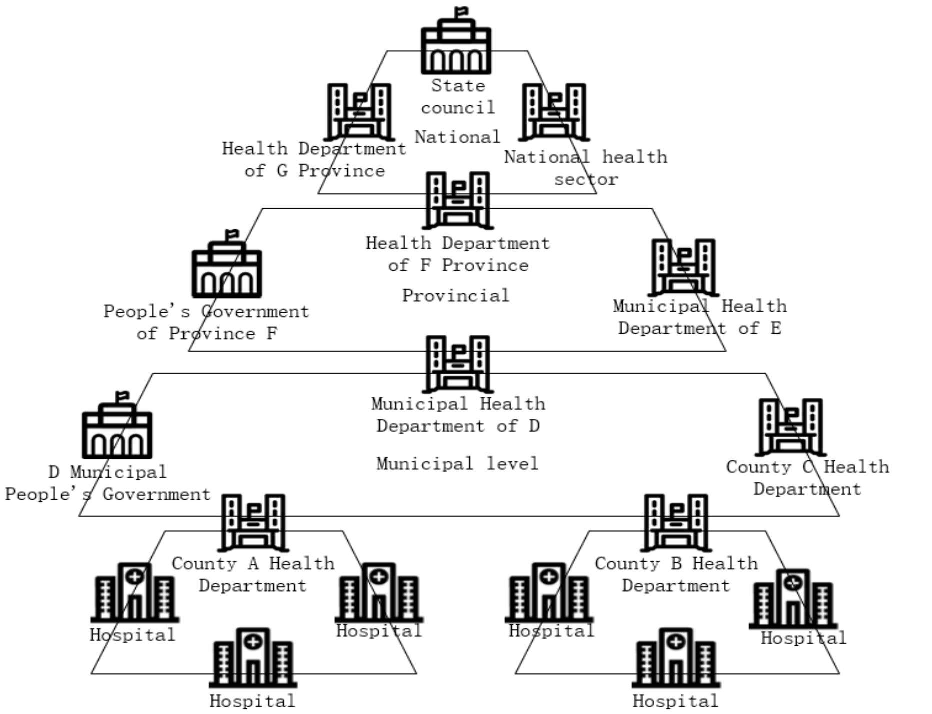

Figure 5. National epidemic early warning decision support chain based on side chain technology.

4.2.3. Human-machine hybrid consensus based on smart contract

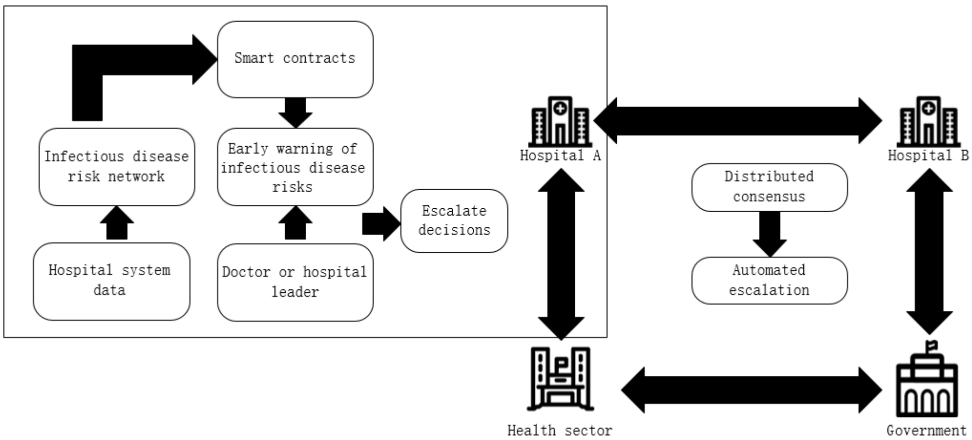

Figure 6. Human-machine hybrid consensus process based on smart contracts.

In the decision-making process for early warning and reporting of infectious illnesses, the conventional reporting paradigm that depends on physicians' diagnoses includes issues such as inadequate personal experience, work omissions, and excessive reporting pressure. If only the algorithm model is utilized to judge, algorithm bias may exist. This research thus presents a hybrid human-machine decision-making system based on smart contracts and consensus procedures. When the danger value hits the threshold, the smart contract-based automated warning is paired with the doctor's evaluation and comments. The institutions on the blockchain establish an agreement on the warning and review process and record it as a responsibility on the blockchain. It may optimize the benefits of human-machine intelligence and consider the timeliness and precision of early warning. **Figure 6** depicts the consensus process based on smart contracts and hybrid human-machine decision-making. First, many medical facilities in the area are operationally and financially based. The infectious illness risk is computed based on the case data. When the regional infectious disease risk surpasses the predetermined level, the smart contract will give an early warning to the physician and the responsible party. The physician will review the caution. system for early detection and correction of deviations. The early warning, review, and reporting procedures must all be published on the blockchain. After each medical institution obtains a distributed agreement on the reporting outcome, it will be communicated automatically to the department at the next level. Also, because the process of early warning, auditing, and reporting is recorded in the blockchain, an unchangeable data record is made. This helps build a system for tracing who is responsible for early warning and auditing of infectious diseases.

5. Conclusion

This paper proposed a novel approach for early warning and detection of infectious diseases based on complex network modeling and blockchain technology. The proposed method leverages hospital case data and operational data to construct an infectious disease risk network, which can identify abnormal group symptoms and medical resource configurations in a timely manner. The use of smart contracts and a hybrid human-machine reporting decision-making technique enables an automated reporting system for regional disease hazards based on distributed consensus. This research contributes to the theoretical system of infectious disease prediction, complex modeling, and consensus algorithms. The proposed method is expected to improve the hospital's ability to detect and respond to emerging infectious diseases, enhance the timeliness and accuracy of infectious disease reporting, and promote joint epidemic reporting and transparent decision-making. Overall, this work has significant implications for public health and can help prevent and control infectious disease outbreaks more effectively.

Author contributions

Conceptualization, YW and MZ; methodology, YC and JM; software, YC and JM; validation, JMZ, YC and JM; formal analysis, YC and JM; investigation, YW and MZ; resources, YW and MZ; data curation, YW and MZ; writing—original draft preparation, YC and JM; writing—review and editing, YC and JM; visualization, JMZ, YC and JM; supervision, JMZ; project administration, JMZ; funding acquisition, JMZ. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Xiong L, Hu P, Wang H. Establishment of epidemic early warning index system and optimization of infectious disease model: Analysis on monitoring data of public health emergencies. International Journal of Disaster Risk Reduction 2021; 65: 102547. doi: 10.1016/j.ijdrr.2021.102547
2. Goh TN. Six sigma in industry: Some observations after twenty-five years. Quality and Reliability Engineering International 2011; 27(2): 221–227. doi: 10.1002/qre.1093
3. Tang W, Liao H, Marley G, et al. The changing patterns of coronavirus disease 2019 (COVID-19) in China: A tempogeographic analysis of the severe acute respiratory syndrome coronavirus 2 epidemic. Clinical Infectious Diseases 2020; 71(15): 818–824. doi: 10.1093/cid/ciaa423
4. Henrickson M. Kiwis and COVID-19: The Aotearoa New Zealand response to the global pandemic. The International Journal of Community and Social Development 2020; 2(2): 121–133. doi: 10.1177/2516602620932558
5. Lan Y, Zhou D, Zhang H, et al. Development of early warning models. In: Yang W (editor). Early Warning for Infectious Disease Outbreak. Academic Press; 2017. pp. 35–74.
6. Yang W. Early Warning of Infectious Disease Theory and Practice. Academic Press; 2000.
7. Buckeridge DL, Okhmatovskaia A, Tu S, et al. Understanding detection performance in public health surveillance: Modeling aberrancy-detection algorithms. Journal of the American Medical Informatics Association 2008; 15(6): 760–769. doi: 10.1197/jamia.M2799

8. Jia P, Yang S. Early warning of epidemics: Towards a national intelligent syndromic surveillance system (NISSS) in China. British Medical Journal Global Health 2020; 5(10): e002925. doi: 10.1136/bmjgh-2020-002925
9. Sellick JA. The use of statistical process control charts in hospital epidemiology. Infection Control and Hospital Epidemiology 1993; 14(11): 649–656. doi: 10.1086/646659
10. Baker AW, Nehls N, Ilieş I, et al. Use of optimised dual statistical process control charts for early detection of surgical site infection outbreaks. British Medical Journal Quality and Safety 2020; 29(6): 517–520. doi: 10.1136/bmjqs-2019-010586
11. Howard L. Statistical process control: A quantitative approach to ensuring quality. The Admitting Management Journal 1990; 15(3): 6–7.
12. Tao H, Zain JM, Band SB, et al. SDN-assisted technique for traffic control and information execution in vehicular adhoc networks. Computers and Electrical Engineering 2022; 102: 108108. doi: 10.1016/j.compeleceng.2022.108108
13. Hai T, Said NM, Zain JM, et al. ANN usefulness in building enhanced with PCM: Efficacy of PCM installation location. Journal of Building Engineering 2022; 57: 104914. doi: 10.1016/j.jobe.2022.104914
14. Hai T, Abidi A, Zain JM, et al. Assessment of using solar system enhanced with MWCNT in PCM-enhanced building to decrease thermal energy usage in ejector cooling system. Journal of Building Engineering 2022; 55: 104697. doi: 10.1016/j.jobe.2022.104697
15. Hai T, Abidi A, Wang L, et al. Thermal analysis of building benefits from PCM and heat recovery-installing PCM to boost energy consumption reduction. Journal of Building Engineering 2022; 58: 104982. doi: 10.1016/j.jobe.2022.104982
16. Hai T, El-Shafay AS, Zain JM, et al. The effect of triangular phase change material rods in the air conditioning duct on the amount of energy required for a residential building. Journal of Building Engineering 2022; 52: 104330. doi: 10.1016/j.jobe.2022.104330
17. Zhang H, Li Z, Lai S, et al. Evaluation of the performance of a dengue outbreak detection tool for China. PLoS One 2014; 9(8): e106144. doi: 10.1371/journal.pone.0106144
18. Hutwagner L, Thompson W, Seeman GM, et al. The bioterrorism preparedness and response early aberration reporting system (EARS). Journal of Urban Health 2003; 80(1): i89–i96. doi: 10.1007/PL00022319
19. Li Z, Lai S, Zhang H, et al. Hand, foot and mouth disease in China: Evaluating an automated system for the detection of outbreaks. Bulletin of the World Health Organization 2014; 92: 656–663. doi: 10.2471/BLT.13.130666
20. Unkel S, Farrington CP, Garthwaite PH, et al. Statistical methods for the prospective detection of infectious disease outbreaks: A review. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2012; 175(1): 49–82. doi: 10.1111/j.1467-985X.2011.00714.x
21. Takahashi K, Shimadzu H. Detecting multiple spatial disease clusters: Information criterion and scan statistic approach. International Journal of Health Geographics 2020; 19(1): 1–11. doi: 10.1186/s12942-020-00228-y
22. Shariati M, Mesgari T, Kasraee M, et al. Spatiotemporal analysis and hotspots detection of COVID-19 using geographic information system (March and April, 2020). Journal of Environmental Health Science and Engineering 2020; 18(2): 1499–1507. doi: 10.1007/s40201-020-00565-x
23. Saffary T, Adegboye OA, Gayawan E, et al. Analysis of COVID-19 cases' spatial dependence in US counties reveals health inequalities. Frontiers in Public Health 2020; 8: 579190. doi: 10.3389/fpubh.2020.579190
24. Kulldorff M. Spatial scan statistics: Models, calculations, and applications. In: Glaz J, Balakrishnan N (editors). Scan Statistics and Applications. Birkhäuser; 1999. pp. 303–322.
25. Yang W, Li Z, Lan Y, et al. A nationwide web-based automated system for outbreak early detection and rapid response in China. Western Pacific Surveillance and Response Journal 2011; 2(1): 10–15. doi:10.5365/WPSAR.2010.1.1.009
26. Kulldorff M, Heffernan R, Hartman J, et al. A space-time permutation scan statistic for outbreak disease detection. Public Library of Science Medicine 2005; 2(3): e59. doi: 10.1371/journal.pmed.0020059
27. Hohl A, Delmelle EM, Desjardins MR, et al. Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. Spatial and Spatio–temporal Epidemiology 2020; 34: 100354. doi: 10.1016/j.sste.2020.100354
28. Takahashi K, Kulldorff M, Tango T, et al. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. International Journal of Health Geographics 2008; 7(1): 1–14. doi: 10.1186/1476-072X-7-14
29. Kraemer MU, Reiner RC, Brady OJ, et al. Past and future spread of the arbovirus vectors Aedes aegypti and Aedes albopictus. Nature Microbiology 2019; 4(5): 854–863. doi: 10.1038/s41564-019-0376-y
30. Odhiambo JN, Kalinda C, Macharia PM, et al. Spatial and spatio-temporal methods for mapping malaria risk: A systematic review. British Medical Journal Global Health 2020; 5(10): e002919. doi: 10.1136/bmjgh-2020-002919
31. Siettos CI, Russo L. Mathematical modeling of infectious disease dynamics. Virulence 2013; 4(4): 295–306. doi: 10.4161/viru.24041
32. Yang W, Lan Y, Li Z. Review and prospects of infectious disease early warning research. Chinese Journal of Preventive Medicine 2014; 48(4): 244–247.

33. Wu J, Wang J, Nicholas S, et al. Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations. Journal of Medical Internet Research 2020: 22(10): e21980. doi: 10.2196/21980
34. Liu L, Luan RS, Yin F, et al. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. Epidemiology and Infection 2016; 144(1): 144–151. doi: 10.1017/S0950268815001144
35. Yang E, Park HW, Choi YH, et al. A simulation-based study on the comparison of statistical and time series forecasting methods for early detection of infectious disease outbreaks. International Journal of Environmental Research and Public Health 2018; 15(5): 966. doi: 10.3390/ijerph15050966
36. Carneiro HA, Mylonakis E. Google trends: A web-based tool for real-time surveillance of disease outbreaks. Clinical Infectious Diseases 2009; 49(10): 1557–1564. doi: 10.1086/630200
37. Jia W, Wan Y, Li Y, et al. Integrating multiple data sources and learning models to predict infectious diseases in China. American Medical Informatics Association Joint Summits on Translational Science Proceedings 2019; 2019: 680–685.
38. Ceballos-Arroyo AM, Maldonado-Perez D, Mesa-Yepes H, et al. Towards a machine learning-based approach to forecasting Dengue virus outbreaks in Colombian cities: A case-study: Medellin, Antioquia. In: 15th International Symposium on Medical Information Processing and Analysis; 6–8 November 2019; Medellín, Colombia. pp. 349–359.
39. Hu S, Li J, Liu F, et al. Establishment and response of SARS surveillance and precaution system for early-stage symptoms in Hunan province. Disease Surveillance 2005; 20(7): 344–347. doi: 10.3784/j.issn.1003-9961.2005.7.344
40. Oeschger TM, McCloskey DS, Buchmann RM, et al. Early warning diagnostics for emerging infectious diseases in developing into late-stage pandemics. Accounts of Chemical Research 2021; 54(19): 3656–3666. doi: 10.1021/acs.accounts.1c00383
41. Hai T, Alsharif S, Dhahad HA, et al. The evolutionary artificial intelligence-based algorithm to find the minimum GHG emission via the integrated energy system using the MSW as fuel in a waste heat recovery plant. Sustainable Energy Technologies and Assessments 2022; 53: 102531. doi: 10.1016/j.jobe.2022.104330
42. Hai T, Dhahad HA, Attia EA, et al. Proposal 3E analysis and multi-objective optimization of a new biomass-based energy system based on the organic cycle and ejector for the generation of sustainable power, heat, and cold. Sustainable Energy Technologies and Assessments 2022; 53: 102551. doi: 10.1016/j.seta.2022.102551
43. Hai T, Zhou J, Srividhya SR, et al. BVFLEMR: An integrated federated learning and blockchain technology for cloud-based medical records recommendation system. Journal of Cloud Computing 2022; 11(1): 22. doi: 10.21203/rs.3.rs-1750276/v1
44. Peters KE, Walters CC, Moldowan JM. The Biomarker Guide: Volume 1, Biomarkers and Isotopes in the Environment and Human History, 2nd ed. Cambridge University Press; 2007.
45. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. Physical Review Letters 2001; 86(14): 3200. doi: 10.1103/PhysRevLett.86.3200
46. Moreno Y, Pastor-Satorras R, Vespignani A. Epidemic outbreaks in complex heterogeneous networks. European Physical Journal B 2002; 26: 521–529. doi: 10.1140/epjb/e20020122
47. Li CH, Tsai CC, Yang SY. Analysis of epidemic spreading of an SIRS model in complex heterogeneous networks. Communications in Nonlinear Science and Numerical Simulation 2014; 19(4): 1042–1054. doi: 10.1016/j.cnsns.2013.08.033
48. Yang Z, Zhang J, Gao S, et al. Complex contact network of patients at the beginning of an epidemic outbreak: An analysis based on 1218 COVID-19 cases in China. International Journal of Environmental Research and Public Health 2022; 19(2): 689. doi: 10.3390/ijerph19020689
49. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. Public Library of Science One 2010; 5(9): e12948. doi: 10.1371/journal.pone.0012948
50. Holme P, Saramäki J. Temporal networks. Physics Reports 2012; 519(3): 97–125. doi: 10.1016/j.physrep.2012.03.001
51. Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: Challenges, methods, and future directions. arXiv: 1908.07873. doi: 10. .48550/arXiv.1908.07873
52. Kim H, Park J, Bennis M, et al. Blockchained on-device federated learning. Institute of Electrical and Electronics Engineers Communications Letters 2019; 24(6): 1279–1283. doi: 10.1109/LCOMM.2019.2921755