New theoretical ISM-K2 Bayesian network model for evaluating vaccination effectiveness

Xiaoliang Xie[1,3] · Bingqi Xie[1,4] · Dan Xiong[2] · Muzhou Hou[2] · Jinxia Zuo[1,4] · Guo Wei[5] · Julien Chevallier[6,7]

Abstract

Aiming at the difficulty in obtaining a complete Bayesian network (BN) structure directly through search-scoring algorithms, authors attempted to incorporate expert judgment and historical data to construct an interpretive structural model with an ISM-K2 algorithm for evaluating vaccination effectiveness (VE). By analyzing the influenza vaccine data provided by Hunan Provincial Center for Disease Control and Prevention, risk factors influencing VE in each link in the process of "Transportation—Storage—Distribution—Inoculation" were systematically investigated. Subsequently, an evaluation index system of VE and an ISM-K2 BN model were developed. Findings include: (1) The comprehensive quality of the staff handling vaccines has a significant impact on VE; (2) Predictive inference and diagnostic reasoning through the ISM-K2 BN model are stable, effective, and highly interpretable, and consequently, the post-production supervision of vaccines is enhanced. The study provides a theoretical basis for evaluating VE and a scientific tool for tracking the responsibility of adverse events of ineffective vaccines, which has the value of promotion in improving VE and reducing the transmission rate of infectious diseases.

✉ Muzhou Hou
houmuzhou@sina.com

1 School of Mathematics and Statistics, Hunan University of Technology and Busin Ess, Changsha 410205, China

2 School of Mathematics and Statistics, Central South University, Changsha 410083, China

3 Key Laboratory of Hunan Province for Statistical Learning and Intelligent Computation, Hunan University of Technology and Business, Changsha 410205, Hunan, China

4 Institute of Big Data and Internet Innovation, Hunan University of Technology and Business, Changsha 410205, China

5 Department of Mathematics and Computer Science, University of North Carolina at Pembroke, Pembroke, NC 28372, USA

6 IPAG Business School (IPAG Lab), 184 boulevard Saint-Germain, 75006 Paris, France

7 University Paris 8 (LED), 2 rue de la Liberté, 93526 Saint-Denis, France

1 Introduction

During the COVID-19 pandemic, production economics and operational management are called upon to support decision-making processes. In this burgeoning literature, Ivanov and Dolgui (2020) unveiled unforeseen and unprecedented fragilities in supply chains. They scrutinized the existing operational research studies dealing with disruption propagation and structural dynamics in supply chains. Dasaklis et al. (2012) predicted that outbreaks of epidemics will account for a great number of deaths. They assessed that effective control of an epidemic outbreak calls for a rapid response: "Available resources such as essential medical supplies and well-trained personnel need to be deployed rapidly and to be managed in conjunction with available information and financial resources in order to contain the epidemic before it reaches uncontrollable or disastrous proportions." Their latter prediction became sadly very true in 2020.

Vaccines are vital to human survival, health and development, and are hot topics of common concerns for governments, experts, and scholars around the world (CDC 1999). When infectious diseases strike, in addition to providing timely assistance to infected patients, the development of vaccines has become critically important. From a domestic

perspective in China, the Changchun Changsheng Vaccine Event in 2018 caused a sensation throughout the country; The number of problematic vaccines investigated was as high as 650,000, of which 247,359 have been already vaccinated, and this involved 215,184 children. The novel coronavirus pneumonia (COVID-19), which was noticed in January 2020 in Wuhan, China, as of May 2020, went outbreak and has become a worldwide pandemic, forcing the World Health Organization (WHO) to declare it as a most severe world public health crisis. Public health experts come to a consensus on the urgent needs of vaccine-type preventive biological products to save people's life. Therefore, the development and safety of vaccines have been the primary issue in fighting infectious diseases, preventing sudden health emergencies, and managing sudden life-threatening public health events.

Major international infectious diseases in recent years include Severe Acute Respiratory Syndrome (SARS), Ebola Virus Disease, Middle East Respiratory Syndrome (MERS-CoV), and the most recent Novel Coronavirus Pneumonia (NCP or COVID-19). According to WHO, as of August 16, 2003, the number of patients with SARS worldwide was 8,096, the number of deaths was 774, and the mortality rate was 9.6% (WHO 2020f). From 2012 until January 15, 2020, the total number of laboratory-confirmed MERS infection cases reported globally to WHO was 2,506, with 862 associated deaths (WHO 2020e). The Ebola virus has considerable epidemic potential. For example, the scale of the West African epidemic in 2013–2016 was unprecedented, with more than 28,000 confirmed cases and 11,000 deaths. For the Ebola infectious disease incidents in Africa, as of February 11, 2020, a total of 3432 Ebola cases were reported, including 3309 confirmed cases and 123 suspected cases, of which 2253 died, with a mortality rate of 66% (WHO 2020c).

At present, as of May 19, 2022, COVID-19 in China has caused 1,879,192 confirmed cases and 15,713 deaths (WHO COVID-19 Dashboard), and more severe outbreaks happen worldwide out of control, as detailed in the next paragraph. The main transmission routes are respiratory droplets and contact transmission, which are highly contagious. People of all ages may be infected, mainly adults, and the elderly and frail people seem to be more likely to be infected (WHO 2020b).

According to WHO, as of May 19, 2022, the total number of infected people by COVID-19 is over 520million, and total deaths are over 6 million (6,270,232), still rapidly growing. The breakdown of these numbers by countries are as follows: China (infected count: 1,879,192, Deaths: 15,713), United States (over 81 million, 992,667), India (over 43 million, 524,293), England (over 22 millon, 177,580), Italy (over 17 millon, 165,494), Spain (over 12 millon, 105,642), France (28 millon, 144,302), Brazil (30 millon, 664,987), Iran (over 7 millon, 141,236).

Infectious diseases in China also include hand-foot-mouth disease, H1N1 influenza, and so on. From 2008 to 2009, a total of 1,644,480 cases of hand, foot, and mouth disease were reported in China, 14,977 patients were severe, and 479 people died. Among them, infants and young children were mainly under five years old. The infection routes of hand, foot and mouth diseases (HFMD) include digestive tract, respiratory tract and contact transmission (Zhu et al. 2011). On March 31, 2010, more than 127,000 confirmed cases of H1N1 influenza were reported in 31 provinces across the country, including 800 deaths. On June 11, 2010, WHO stated that 18,156 people worldwide have died since the outbreak of H1N1 influenza a year ago. The transmission of H1N1 influenza is mainly through the respiratory tract, but it can also be transmitted through contact with infected pigs or their feces and the surrounding polluted environment. The susceptible population is young adults aged 20–45 (Vaillant et al. 2009).

The spread of infectious diseases is rapid, broad, and fatal, and it is crucial to carry out epidemic prevention and control promptly. Hence, speeding up the development of vaccines and ensuring safety and VE is the key to confront global virus threats.

At present, relevant studies of vaccination mainly focused on the following aspects: First, the necessity and importance of vaccination. As pointed out by WHO in 2018, current vaccination alone can prevent 2–3 million deaths each year (WHO 2020a). Ventola estimated that the average mortality rate from preventable diseases due to vaccination has decreased by 97.8% (Ventola 2016).

The second is the study of vaccine storage and transportation conditions. The transportation, storage, and other aspects of the vaccine require very strictly restricted temperature and other requirements. When the ambient temperature of a vaccine is outside a moderate range, the VE can be weakened or even completely lost. Patients mistakenly believe that they received a routine vaccine while the vaccine they received is invalid, causing a more significant increase in the probability of catching a disease (WHO 2004). Generally, high temperature has a tremendous negative impact on the effectiveness of almost all vaccines but also is the fact that too low temperature can inactivate freeze-sensitive vaccines such as tetanus toxoid, Haemophilus influenza b (Hib), and inactivated poliovirus (WHO 2020g). Therefore, keeping the vaccines at the optimal temperature range at all times during transportation is critical to maintaining VE (Yakum et al. 2015). In Brazil, the vaccine has to be delivered through the Amazon jungle and a long distance to reach remote communities, along with poor and difficult transportation conditions, causing the overall VE to be reduced (WHO 2020d). According to the statistics, a considerable proportion of vaccines in the course of transportation lose their effectiveness

partially or even wholly because of the exposure to excessively high or sub-zero temperatures (Das et al. 2019). A study (2007–2014) showed that vaccines were exposed to freezing temperatures during storage (29.7%) and transportation (28.1%) in low- and middle-income countries (Hanson et al. 2017). Kinfe et al. believed that, in some developing countries, although vaccination coverage has been increased, due to old-fashioned equipment or newly developed equipment that cannot be upgraded in time (due to the poor economic conditions), the recurrence rate of vaccine-preventable diseases remains high (Kinfe et al. 2019). Kumru found that the requirements for storage and transportation are high (Kumru et al. 2014). Proper, standardized, and scientific storage is the core condition for ensuring vaccine safety.

The third is the study of vaccination effect and influencing factors. As pointed out by Phillips et al. (2017), Brison and LeTallec (2017), in low-income and middle-income countries, many children do not always develop immunity after vaccination. At the same time, inadequate preparation of medical facilities and sanitation facilities and severe cold chain breaks are important factors affecting willingness and vaccination effectiveness. Huang et al., by understanding the positive rate of HBsAb in lymphoma patients after receiving different doses of hepatitis B vaccine, evaluated the immune effect of different doses of Hepatitis B vaccine and analyzed the influencing factors (Zhuang and Wang 2020). Lymphoma patients can get a better immune effect if they receive a high-dose hepatitis B vaccine. Rupprecht and Salahuddin pointed out that the preparation technology of rabies vaccine has been significantly advanced, and the protection rate for rabies vaccine has reached a higher level (about 90%) (Rupprecht and Salahuddin 2019). Still, the occurrence of rabies has not yet been eliminated. Yang et al. investigated the primary immune effect and influencing factors of the hepatitis B vaccine by understanding the response of the 5ug recombinant hepatitis B vaccine in adults (Yang et al. 2016). Head et al. regarded the vaccine efficacy test as the primary indicator for evaluating the immunogenicity of the vaccine and also as a critical item for vaccine quality control (Head et al. 2019).

The fourth is the research on ways and methods to improve the VE. As pointed out by Rogers et al., vaccine staff must understand the current storage of vaccines, be able to handle any errors when occurring, and know how to avoid the mistakes (Rogers et al. 2010). In the case of an event, they must take immediate action for correcting improper storage and handling conditions. Strohfus et al. proposed that after providing immunization education, with comparison pre-test and post-test results, the ability of medical staff to manage vaccine practice has increased by 7.8%, and the efficiency of immunization in the clinics has increased by 10.3% (Strohfus et al. 2017).

The fifth is research on the regulation of vaccination. Tostanoski and Jewell indicated that vaccine cold chain management should be strengthened, supervision should be improved, vaccine management and management staff should be controlled, and vaccine temperature should be maintained at appropriate degrees (Tostanoski and Jewell 2017).

The sixth is the research on VE evaluation methods. Traditional evaluation types include cohort studies, family contact studies, case–control studies, and case studies (Torvaldsen and Mcintyre 2002). Essential issues about VE cover whether an individual has an adverse reaction to the vaccine and the impact of the vaccine's coverage plan on effectiveness, which can be answered based on the observations only after vaccines have been administered. The degree of immunization of a vaccine is affected by many factors, such as whether or not maintained under appropriate storage conditions (such as cold chains, access to vaccines), local health care conditions, vaccine costs, and other non-vaccine-related factors (Weinberg 2010). Falchi adopted the age of the people vaccinated with influenza, whether they have chronic diseases, and the presence of influenza virus as important indicators to evaluate the effectiveness of influenza vaccines but did not study other confounding factors such as health care coverage, gender, previous vaccine vaccination, or smoking (Falchi et al. 2013).

In summary, research contents in the current literature are very active and prosperous and form an essential basis for our study. However, a common hindrance or obstruction lies in the spatial and temporal dispersion of vaccination, complexity of the vaccination staff, and various extreme difficulties in obtaining valid data. Consequently, research on VE still faces the following deficiencies: (1) Qualitative research is predominant, but the quantitative analysis is insufficient; (2) Traditional statistical methods are frequently used, but modern intelligent methods are less employed; and (3) Examining the literature we retrieved, to the best of our knowledge, there are no previous studies on VE analysis using BN. Hence, the purpose of this study is to employ the BN as a tool to investigate the effectiveness and safety of vaccines throughout the transportation process and delivery to the administering sites.

Further on recent developments in the literature of Bayesian networks, we may cite several works. First, Boutselis and McNaught (2019) propose a novel application of Bayesian Networks to spare parts demand forecasting. They found that Expert adjusted forecast outperformed by unsupervised and hybrid Bayesian Networks. Second, Hosseini and Barker (2016) proposed a Bayesian network that effectively models the causal relationships among variables but that has not been used in the context of supplier evaluation and selection to quantify the appropriateness of suppliers across primary, green, and resilience criteria.

More precisely, this paper attempts to utilize a hybrid BN model, along with combining expert questionnaires and historical data, to evaluate the VE. Based on the analysis of risk factors of vaccine adverse events over the years, with the help of using the ISM method (Warfield 2010) and K2 algorithm (Cooper and Herskovits 1992), BN structure learning and parameter learning are performed to establish a mixed Bayesian evaluation of VE network model.

The remainder of the paper is organized as follows. In Sect. 2, a risk assessment index system for VE is constructed, a BN model combined with an interpretative structural model method is proposed, and a BN structure learning is performed. In addition, a BN parameter is estimated based on a MAP algorithm (Heckerman et al. 1995). In Sect. 3, we mainly conduct an empirical study and analyze the results and verify the feasibility of the model through stability, effectiveness analysis, and comparison with other methods. Section 4 draws the conclusion, provides political implications, discusses research limitations, and outlines future research directions. The overall framework of the paper is illustrated in Fig. 1.

2 Theory construction and model building

2.1 Establishing risk index system for VE

This article mainly studies the influenza vaccine, which belongs to the second-class vaccine in China, namely the self-paid vaccine. Figure 2 shows the approval status of influenza vaccines issued by the China National Institutes for Food and Drug Control (NIFDC) in the past 3 years. In 2017–2019, a total of nine vaccine manufacturers produced influenza vaccines. Among them, there was no influenza vaccine listed in Taizhou Jindike in 2017. In 2018, the Shanghai Institute of Biology and Beijing Kexing, Changchun Changsheng, and Taizhou Jindike were not shown in the list of influenza vaccines. In 2019, Shanghai Institute of Biology, Dalian Yalifeng, Tianjin Zhongyi Anke, and Changchun Changsheng were not listed. In 2017, 2018, and 2019, the numbers of influenza vaccines listed in China each were respectively 29.907 million, 16.124 million, and 20.217 million (NIFDC 2020).

According to the reports by the United States Centers for Disease Control and Prevention (CDC), the flu season has caused 22–31 million infections and 12,000 to 30,000 deaths in the US. from 2019 to 2020. China's population is four times that of the United States, and the impact of influenza on the people is even more significant. According to statistics from the US CDC, from 2019 to 2020, American influenza vaccine manufacturers predicted that they would provide up to 162 million to 169 million doses of influenza vaccines to the market. From 2019 to 2020, of the 1.395 billion people in China, only some 20 million doses were administered to a small percentage of the population; In contrast, the US has about 327 million people and 169 million doses were administered. The two countires's vaccination rates were 2% and 51% respectively. Compared to the US, the flu vaccination rate in China has a large gap.
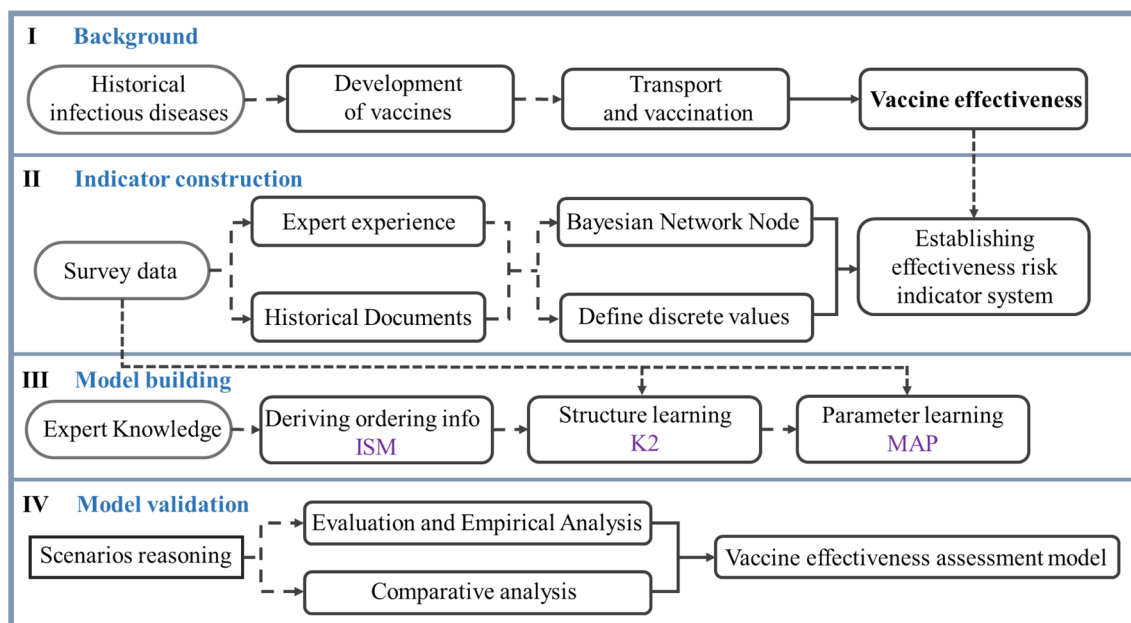

Fig. 1 Framework and flowchart of the paper

Fig. 2 Approval of China's influenza vaccine market. Data Sources: https://www.nifdc.org.cn/nifdc/

Fig. 3 Distribution of influenza vaccines from origins to Hunan Province

Figure 3 shows the distribution of influenza vaccine manufacturers that supply Hunan Province. In recent years, as there are no vaccine manufacturers located in Hunan Province, influenza vaccines have to be transported from various originsthrough a cold chain. Since the temperature of the vaccine has a significant impact on VE, and many factors affect VE during the transportation process, it is vital to conduct a VE risk assessment to ensure VE

Risk assessment indicators for VE are selected by incorporating the following two aspects: (1) Research results from the existing literature, preliminary design indicators from the perspective of cold chain transportation, reports and case studies as well as relevant analyses from the practice, and comprehensively refined risk factors from qualitative and quantitative studies (Head et al. 2019; Phillips et al. 2017; Yang et al. 2016); and (2) Building a risk assessment index system for VE from four dimensions (a through d in Table 1 below), by referring to historical literature and expert opinions from Hunan provincial CDC (Rogers et al. 2010; Tostanoski and Jewell 2017).

This indicator system is developed to aim at the vaccine production workshop, cold chain storage, cold chain transportation, Hunan provincial CDC storage, vaccination site storage, and other links. In this system, the risk factors are initially classified, possible risk factors included in each link are analyzed, and a risk assessment index system for VE is constructed. The system consists of eleven indicators, and the assessment criteria of each indicator are referred to as the opinions of experts from the Hunan provincial CDC. Among them, the evaluation of VE is obtained through testing and calculation, as shown in Table 2. The establishment of this indicator system focuses on external factors, namely from the perspective of the vaccine transport chain. The eleven indicators are set up to refine the risk factors and improve the accuracy of the assessment.

2.2 BN structure learning with hybrid ISM method

BN structure learning can be designed to incorporate expert experience or learn from data, but directly acquiring knowledge from experts can be time-consuming and error-prone. In the case of a large number of variables in a BN, it is often difficult for domain experts to describe the structural dependencies and independence between variables. On the other hand, for areas where human experts have limited knowledge but sufficient cases and sufficient data available, through search-scoring algorithms, the BN structure can be learned from the data. However, it is challenging to automatically discover the structure from the data because there are many possible structures, which would require massive high-quality data for rigorous detection.

As China does not yet have a complete vaccine traceability system, through the cooperation with the Hunan provincial CDC, staff from its subordinate departments have collected influenza data from nine influenza vaccine manufacturers and vaccination sites across the country from 2017 to 2019, including data on vaccine delivery and vaccination. Data on the preconceived risk assessment indicators were collected. Currently, influenza vaccines that have been marketed internationally include influenza inactivated vaccines and live attenuated influenza vaccines. The influenza vaccines that have been sold in China are both trivalent inactivated vaccines and tetravalent inactivated vaccines. Trivalent vaccines are administered for infants over six months, children, and adults to prevent influenza caused by type A H1N1, type A H3N2, and type B Victoria. The tetravalent component is one more than the trivalent vaccine, that is, the trivalent vaccine prevents influenza caused by the Yamagata type B, and the tetravalent is used for children over three years old and adults. In this experiment, the composition of the vaccine is not involved, so no distinction is made between trivalent and tetravalent vaccines, and they are collectively referred to as influenza vaccines.

These data include:

1. Ranking of vaccine manufacturers' cities by GDP in all cities of their province (last year).
2. Comprehensive quality of the refrigerated truck driver.
3. Refrigerator compartment temperature records of production and transportation link.
4. Storage room temperature records of the Hunan provincial CDC storage link.
5. Comprehensive quality of storage guards of the Hunan provincial CDC storage link.
6. Refrigerator compartment temperature records of vaccine delivery link.
7. Comprehensive quality of vaccination nurses at vaccination sites.
8. Temperature record in the storage room of vaccination sites.
9. Ranking of vaccination sites' cities by GDP in all cities of their province (Last year).
10. Temperature on the day of inoculation.
11. Evaluation of VE (Zapata 2017).

In the above list, data of the first 11 nodes are obtained through field surveys, and the last node, the assessment of VE, needs to be tested. The usual methods for testing VE include the negative test design method (Foppa et al. 2013; Jackson and Nelson 2013) and the antibody titer test method is commonly used to evaluate the potency and VE (Ganz et al. 1978). This study selected the antibody titer test method. The explanation of how to evaluate the effectiveness after vaccination is as follows:

One vaccine is randomly selected from each batch of vaccines as an experimental sample. The sampling procedure is

Table 1 Construction of VE risk warning indicators

Link	Index	Symbol	Interpretation of indicator discretization
Production and transportation link	Ranking of vaccine manufacturers' cities by GDP in all cities of their province (Last year)	a1	The city accounts for the top 30% of the province's GDP rankings: 1 The city accounts for the middle 30% of the province's GDP rankings: 2 The city accounts for the last 40% of the province's GDP rankings: 3
	Comprehensive quality of refrigerated truck driver	a2	Drivers with a driving age of 8 years or above or a college degree: 1 Drivers with a driving age of 5 years or above or a high school degree: 2 Drivers with a driving age of 3 years or above or a junior high school education: 3
	Refrigerator compartment temperature records	a3	Temperature records are between 2–8 degrees Celsius: 1 Temperature records exceeding 8 degrees Celsius and not exceeding 0.5 h: 2 Temperature record exceeds 8 degrees Celsius for more than 0.5 h: 3
Hunan provincial CDC storage link	Storage room temperature records	b1	Temperature records are between 2–8 degrees Celsius: 1 Temperature records exceeding 8 degrees Celsius and not exceeding 0.5 h: 2 Temperature record exceeds 8 degrees Celsius for more than 0.5 h: 3
	Comprehensive quality of storage guards	b2	Staff's education is junior college or above: 1 Staff's education is high school and above: 2 Staff's education is junior high school and above: 3
Vaccine delivery link	Personal qualities of delivery staff	c1	Staff's education is junior college or above: 1 Staff's education is high school and above: 2 Staff's education is junior high school and above: 3
	Refrigerator compartment temperature records	c2	Temperature records are between 2 and 8 °C: 1 Temperature records exceeding 8 °C and not exceeding 0.5 h: 2 Temperature record exceeds 8 °C for more than 0.5 h: 3
Storage at the vaccination site link	Comprehensive quality of vaccination nurses at vaccination sites	d1	The job title of the staff is Nurse-in-charge and above: 1 The job title of the staff is Nurse Practitioner: 2 The job title of the staff is Nurse: 3
	Temperature record in the storage room of vaccination sites	d2	Temperature records are between 2–8 degrees Celsius: 1 Temperature records exceeding 8 °C and not exceeding 0.5 h: 2 Temperature record exceeds 8 °C for more than 0.5 h: 3
	Ranking of vaccination sites' cities by GDP in all cities of their province (Last year)	d3	The city accounts for the top 30% of the province's GDP rankings: 1 The city accounts for the middle 30% of the province's GDP rankings: 2 The city accounts for the last 40% of the province's GDP rankings: 3
	Temperature on the day of inoculation	d4	Daily maximum temperature is 15 ° and below: 1 Daily maximum temperature is above 15 degrees and below 30 degrees: 2 Daily maximum temperature is above 30 degrees: 3

Table 1 (continued)

Link	Index	Symbol	Interpretation of indicator discretization
Evaluation of vaccination effectiveness	Vaccination effectiveness	e	Judged by experts of Hunan provincial CDC
			One month after inoculation, the antibody titer is qualified: 1
			Low antibody titer after one month of inoculation: 2
			Test antibody titer is zero after one month of inoculation: 3

Table 2 Valuation of V.E

Availability	Grade	Indicators	Remarks
Effective	1	No warning	One month after inoculation, the antibody titer is qualified
Worse	2	Light warning	Low antibody titer after one month of inoculation
Invalid	3	Heavy warning	Test antibody titer is zero after one month of inoculation

used to ensure that the samples are representative and with required sample sizes for the study (estimated to be 400). We selected some participants from those who came to the community hospital to receive the flu vaccine and who also were willing to cooperate with our research and sign a consent form, stating that "after seeking the consent of the vaccine vaccinators, they are required to return to the community hospital for an antibody titer test one month after the vaccination." Since it takes 2–4 weeks for the influenza vaccine to produce protective antibodies, performing an antibody titer test one month after vaccination is the best time. Results of the antibody titer test are closely related to VE

In experimental research, there are often some "non-processing factors" or confounding factors, which will affect the research results, such as gender and age. Confounding factors are among the essential concepts in observational and controlled studies (Pearl et al. 1999). In epidemiological observation research, confounding is a core concept in causal inference, and it is the leading cause of false correlation (no causality but still related). It must be controlled during causal inference. The randomized experiment is the most straightforward and effective method for evaluating causality. Through randomization, the distribution of known and unknown confounding factors can be balanced in the training group and the test group, which means that the randomized experiment is adopted to eliminate known and unknown confounding factors' results in average causal effects (Pourhoseingholi et al. 2012).

A total of 402 sets of experimental data were collected in this study. First, the data was preprocessed to check for outliers. In general, we tend to ignore outliers when building models. This is not a wise practice, as outliers shift data and reduce accuracy. An outlier is an observation that goes well beyond the overall pattern in the sample. The causes of outliers can be divided into two categories: human error
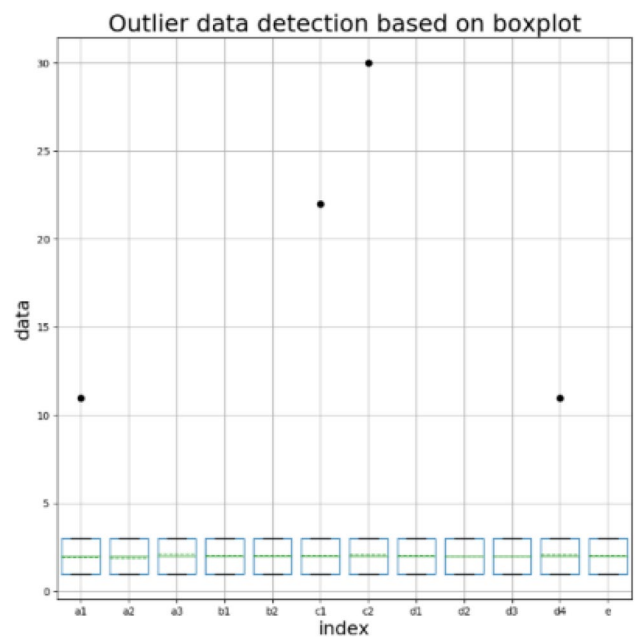
Fig. 4 Outlier data detection based on a boxplot

and natural error. Human errors, such as errors caused during data collection, recording, or entry, can cause outliers in the data. Because the data in this study were collected through professional surveys and artificially discretized, human errors may exist. Sorting the data table is a simple but effective way to highlight outliers. Just sort the data table for each variable and look for unusually high or low values. As shown in Fig. 4, the sample data set has been sorted in ascending order, and the highest value is significantly different from the other values. Although this method cannot quantify, the abnormality of outliers, abnormally high or low values is easy to find. Besides, visualize the data through a

box plot, which can also detect outliers intuitively, as shown in Fig. 4. There are four abnormal points in the figure, the values are: a1 = 11, c1 = 22, c2 = 30, and d4 = 11. These outliers correspond to errors caused by the researcher's input process. Deleting abnormal observations can also be done by using other statistical models to predict the "true" observations occupied by outliers and then adopt the predicted values of these true values in the statistical analysis.

In the process of preprocessing the data, it was found that there were several missing data. Besides, the abnormal values were treated as missing data and were made up through missing data techniques and the Rubin's EM algorithm. In the end, there were a total of six missing data. Therefore, an EM algorithm was employed to handle the missing data. The EM algorithm is iterative. Each iteration can ensure that the likelihood function value increases and converges to a local maximum. The basic principles and implementation steps of the EM algorithm were analyzed, the missing values are interpolated, and a complete experimental data set was established at the end. Compared with directly excluding missing data for experiments, the introduction of the EM algorithm to fill missing data does not need to discard any data, making the data more comprehensive and the accuracy of the experimental results higher.

Selecting the "train_test_split" method of the "Scikit-Learn library" of Python software, we randomly divided the 402 sets of data into training data and test data: Using the cross-validation method to determine the ratio of the training set to the test set to approximately 3: 1, resulting in 302 cases in the set training data and 100 cases in the set of test data. The cross-validation accuracy of the BN model on the training set is shown in Table 3. "Train_test_split" is a commonly used function in cross-validation, and its function is to randomly select train data and test data from the sample in proportion. Then the feature extraction and discretization of the data were carried out, and finally, the data is used for structure learning and parameter learning. Matlab software was used to test the stability of the model using the ten-fold cross-validation method, as shown in Table 3.

2.2.1 Step 1: Domain experts establish pairwise interaction/influence relationships between variables

The 302 sets of training set data were used as the data for structure learning and parameter learning. The K2 algorithm is selected to build the network structure. Still, because the K2 algorithm can be susceptible to the order of variables in the initial setting, the ISM method is adopted to obtain a reasonable initial node ordering for the K2 algorithm. An influence graph is deduced, which represents the direct relationship between any two nodes. To prevent experts from judging the unlikely relationship, the following assumptions are made: (1) Factors in each link affect each other;

Table 3 Accuracy and mean of ten-fold cross-validation method

Accuracy/%										Mean/%
87.9	92.3	91	92	95	89.5	92	91	87.4	88	90.6

(2) Factory shipment may directly affect Hunan provincial CDC storage and VE; (3) Hunan provincial CDC storage Links may directly affect vaccine delivery, and VE; (4) Vaccine distribution links may directly affect vaccination sites, and VE; (5) Vaccine site storage links may directly affect VE After designing the list of possible relationships in advance, the variable relationships were reduced to 52 relationships, as shown in Fig. 5, where "3×2" means that there are 6 (=3×2) one-to-one direct relationships between the three factors, and similar for notation "2×1" and so on. Two groups of domain experts were invited to judge these 52 relationships, discuss the differences and reach consensus, and build a Structural Self-Interaction Matrix (SSIM) to capture the relationship between variables.

2.2.2 Step 2: Establish a structural self-interaction matrix $SSIM = [r_{ij}]_{n \times n}$, and for each pair of variables $K \in (1, \dots, n)$ and X_j in the set $K \in (1, \dots, n)$, express the direct relationship between them by $K \in (1, \dots, n)$

$$r_{ij} = \begin{cases} 1 & X_i \text{ and } X_j \text{ are directly related} \\ 0 & X_i \text{ and } X_j \text{ are not directly related} \end{cases} \quad (1)$$

Therefore, based on the judgment of domain experts, the following SSIM is established.

2.2.3 Step 3: Compute the reachability matrix M by using the SSIM

From the matrix M, one can determine if a connection path between two variables exists. M is computed by $K \in (1, \dots, n)$, where $1 \leq k \leq n$ represents the length of the longest path of the $SSIM$.

2.2.4 Step 4: Divide M into different hierarchical structure influence diagrams

To decompose the reachability matrix M, we first defined the reachability set $R(X_i)$, containing the variables that can be reached from X_i (including X_i), the prior set $Q(X_i)$ consisting of the variables from which X_i can be reached (including X_i), and the reachable set and the preceding intersection set $A = R(X_i) \cap Q(X_i)$. The level extraction is performed according to the condition $R(Xi) = A$. The top variables are extracted first, which is the final goal of the system. After separating the top element, the other elements are extracted in turn until the bottom element is extracted. All variables are divided into nine levels in order:

$L_1 = \{e\}, L_2 = \{d_4\}, L_3 = \{d_3\}, L_4 = \{c_2, d_2\}, L_5 = \{d_1\}, L_6 = \{b_2, c_1\}, L_7 = \{b_1\}, L_8 = \{a_2, a_3\}, L_9 = \{a_1\}.$

2.2.5 Step 5: Draw a directed acyclic graph based on the direct influence relationship

The formed ISM is shown in Fig. 6a and from which the initial ordering of the BN nodes can be obtained:

1. Ranking of vaccine manufacturers' cities by GDP in all cities of their province (Last year) a1.
2. Comprehensive quality of refrigerated truck driver a2 and Refrigerated truck temperature a3.
3. Storage room temperature records of the Hunan provincial CDC storage link b1.
4. Comprehensive quality of storage guards b2 and Personal quality of distribution staff of the vaccine distribution link c1.
5. Comprehensive quality of vaccination nurses at the storage point of the vaccination site d1.

$$SSIM = \begin{array}{c} \\ a1 \\ a2 \\ a3 \\ b1 \\ b2 \\ c1 \\ c2 \\ d1 \\ d2 \\ d3 \\ d4 \\ e \end{array} \begin{array}{c} a1\ a2\ a3\ b1\ b1\ c1\ c2\ d1\ d2\ d3\ d4\ e \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$M = \begin{array}{c} \\ a1 \\ a2 \\ a3 \\ b1 \\ b2 \\ c1 \\ c2 \\ d1 \\ d2 \\ d3 \\ d4 \\ e \end{array} \begin{array}{c} a1\ a2\ a3\ b1\ b1\ c1\ c2\ d1\ d2\ d3\ d4\ e \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$
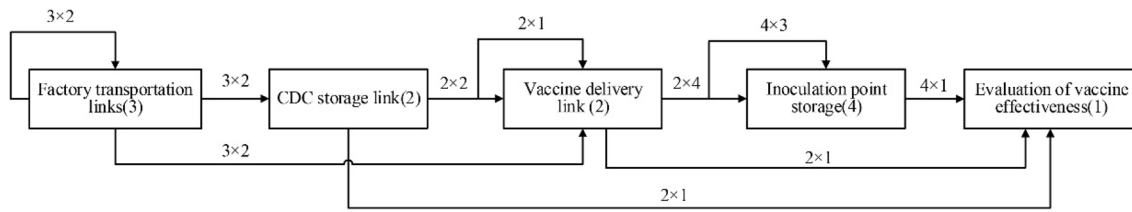
Fig. 5 Providing the relationship between reference variables

6. Temperature of the refrigerated truck in the vaccine distribution process c2 and Storage room temperature of the inoculation site d2.
7. Ranking of vaccination sites' cities by GDP in all cities of their province (Last year) d3.
8. Temperature on the day of the inoculation d4.
9. VE e.

The maximum number of parent nodes of each node is 11, which is the maximum number of parent nodes of the lowest node.

2.2.6 Step 6: Learn BN structure from data

The ordering information of the nodes is used as the input of the initial sequence in the K2 algorithm. The Matlab BNT toolbox is used to learn the BN structure. The final ISM-K2 BN structure is obtained by learning the 302-training data, as shown in Fig. 6b. Observing Fig. 6a and b, a new arrow appears in (b), which also indicates that there is a difference between the expert experience and the actual case. The initial structure can be supplemented and improved

through sufficient case data learning. Compared with the manual construction of BN that rely solely on the opinions of consulting experts, the ISM-K2 structure learning method combines expert experience and historical data information, adding a lot of supplementary information to the final network model, which helps to improve the accuracy of the model reasoning.

2.3 BN parameter learning based on map algorithm

Based on the network structure in Fig. 6b and the 302 training cases (represented by D_r), the MAP algorithm (Zhou et al. 2016) in software GeNIe 2.4 was used to perform parameter learning for determining the following conditional probability: θ_{ijk}, where $Pa(X_i)$ represents the parent node of node i, and θ_{ijk} denotes the conditional probability given that X_i takes the kth value X_i^k and $Pa(X_i)$ takes the jth state $Pa(X_i)^j$. Using the training set D_e, the valuation θ_{ijk} of θ_{ijk} can be calculated by the following formula:

Fig. 6 Risk of vaccination: **a** ISM Model for VE Risk, **b** BN Structure for VE Risk
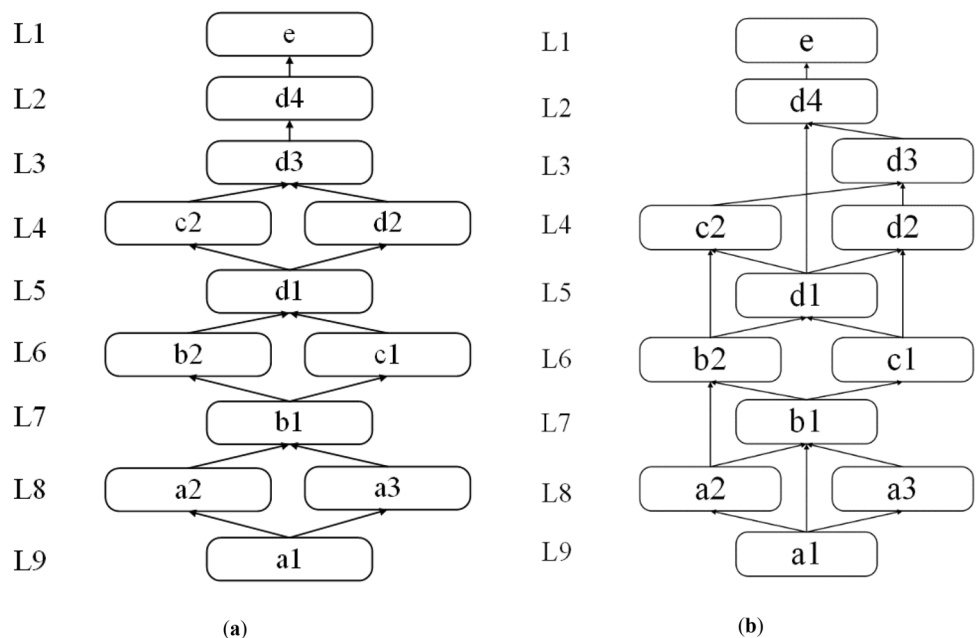
$$\theta_{ijk}^* = \frac{N_{ijk} + \alpha_{ijk}}{N_{ij} + \alpha_{ij}}, \tag{2}$$

where θ_{ijk} counts cases, θ_{ijk} in D_e takes the kth value and its parent node takes the jth value, that is: $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$.

Let $\alpha_{ij} = \sum_{k=1}^{r_i} \alpha_{ijk}$, where θ_{ijk} is a hyper-parameter reflecting the expert's guess on the virtual data count of the parameter θ_{ijk}. If not, enough judgment is obtained from
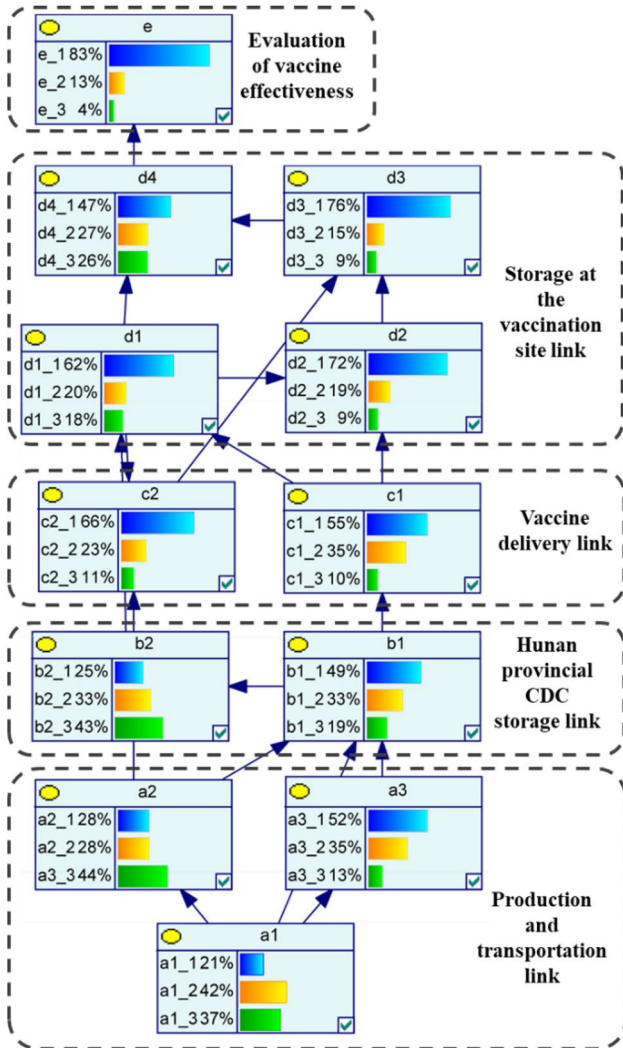
experts in related fields, $\alpha_{ijk} = 1$ or $\alpha_{ijk} = \frac{1}{r_i q_i}$ is generally used [35], where r_i represents the number of values of X_i and q_i counts the parent nodes of X_i. Because of the lack of inputs from experts, $\alpha_{ijk} = \frac{1}{r_i q_i}$ is taken in this study.

The final BN model is given in Fig. 7. "d1_1, c1_1" in Fig. 7 refer to "$d1_1$, $c1_1$" in Table 3, and so on. Each node stores a conditional probability table (CPT). Due to the inconvenience of displaying multiple nodes in the structure, only the CPT of temperature d2 in the storage room at the inoculation site is shown (Table 4).

Among nodes, if the score of comprehensive quality d1 of the nurses at the inoculation site is 2, and the score of personal quality c1 of the distribution personnel at the distribution link is 1, then the probability that the storage room temperature at the inoculation site is qualified is 0.82.

3 Analysis of empirical results and discussions

The test set of 100 cases was applied to verify the model. Factors other than the storage room temperature record (d2) and VE (e) were regarded as known evidence, and they were used as test results. Then a consistency analysis between the prediction results of the BN model from these two nodes and the data in the records is performed. The confusion matrix is shown in Fig. 8a and b summarizes the difference between the predictions from the BN model and the data recorded in the test set. The elements on the diagonal line indicate the number of cases in which the forecast is consistent with the records, and the other elements tell the number of inconsistent cases. The light-yellow color shows a higher degree of difference. In general, 90% and 92% of the predicted data of the two nodes are consistent with the records, and most of the inconsistencies are similar, so they can be considered to be partially consistent.

3.1 Stability analysis

BP neural network is a commonly used algorithm in the field of security risk early warning and assessment. Still, it is susceptible to the initial weight, causing the results of multiple pieces of training not stable. In contrast, the BN has exceptionally high stability. Once the network structure and

Fig. 7 BN for VE Risk

Table 4 CPT of storage room temperature at the inoculation site

d1	$d1_1$			$d1_2$			$d1_3$		
c1	$c1_1$	$c1_2$	$c1_3$	$c1_1$	$c1_2$	$c1_3$	$c1_1$	$c1_2$	$c1_3$
$d2_1$	0.84	0.72	0.59	0.82	0.59	0.48	0.58	0.47	0.38
$d2_2$	0.12	0.23	0.09	0.15	0.32	0.34	0.34	0.23	0.34
$d2_3$	0.04	0.05	0.32	0.03	0.09	0.18	0.08	0.3	0.28

Fig. 8 These are the two figures of the confusion matrix: **a** confusion matrix for temperature record (d2) of inoculation point storage room; **b** confusion matrix for VE (e)

(a)

(b)
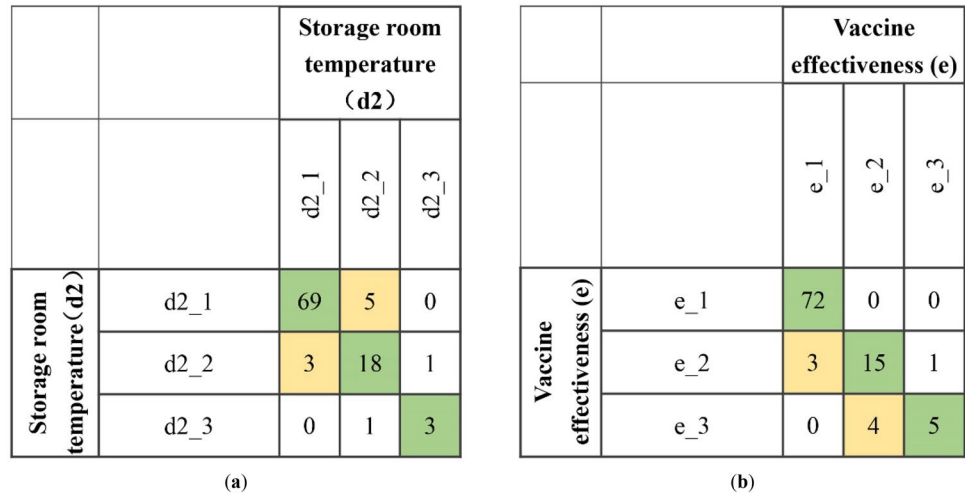
Table 5 Comparison of experimental results between BN and B.P. neural network

Accuracy /%	Number of experiments								Average value /%
	1	2	3	4	5	6	7	8	
BN	90	90	90	90	90	90	90	90	90
BP	91.1	86.7	87.8	89.6	90.6	89.8	90.1	93.5	89.9 ± 3.2

CPT are determined, the operation results will not change. The experiment was repeated eight times, and the vaccine efficacy e was taken as a prediction object. The experimental results are shown in Table 5.

In Table 5, the average accuracy rates of the BP neural network and BN are 89.9% and 90%, respectively. Although the difference is only 0.1%, the results of the eight experiments using the BP neural network method show instability, that is, the accuracy rate of each experiment is different. The highest accuracy is 93.5%, and the lowest is 86.7% giving a big difference between the levels of accuracy. Therefore, after analyzing the stability of the model, we found that the stability of the BN is significantly better than that of the B.P. neural network.

3.2 Effectiveness analysis

To evaluate the ISM-K2 structure learning method in the developed the BN in this study, Firstly, the hill climbing algorithm was compared with the ISM method in structure learning and it was found that the Bayesian network using ISM was more accurate. Then the maximum likelihood estimation (MLE) method was compared with the MAP method in parametric learning, and it was seen that the MAP algorithm yielded conclusions that were closer to reality. Two other models were compared: a BN based on ISM structure learning only and a BN for structural learning using the GA-K2 algorithm. The results of the two methods are shown in Fig. 9. Among them, compared with the ISM-K2

BN, the ISM BN lacks multiple directional connections. That is, the network information is missing. The GA-K2 BN uses a genetic algorithm as a fitness function to obtain the initial order of the nodes and then employs the K2 algorithm structure to learn the BN structure. Therefore, this approach can lead to many abnormal and counterintuitive relationships, such as the arrow from the comprehensive quality of refrigerated truck driver (a2) to the temperature on the day of inoculation (d4). Five experiments were conducted, and the inferred objects were the personal qualities of the distribution personnel (c1), the temperature record of the storage room at the inoculation site (d2), the GDP ranking of the city where the inoculation site was located (d3), the temperature on the day of inoculation (d4), and the temperature record of refrigerated truck (c2). The experimental results are shown in Table 6.

The ISM BN model has a relatively good prediction effect on the GDP ranking of the city where the inoculation site is located (d3). The overall prediction effect of the GA-K2 BN model is better than that of the ISM model. However, compared with the ISM-K2 BN method used in this study, their accuracy rates are relatively low, and the structure of the GA-K2 BN model also has an unconventional directing relationship.

Therefore, the model produced by the ISM-K2 BN has better training and reasoning performance. This method is suitable for BN modeling with relatively fewer data and relatively little domain expert knowledge.
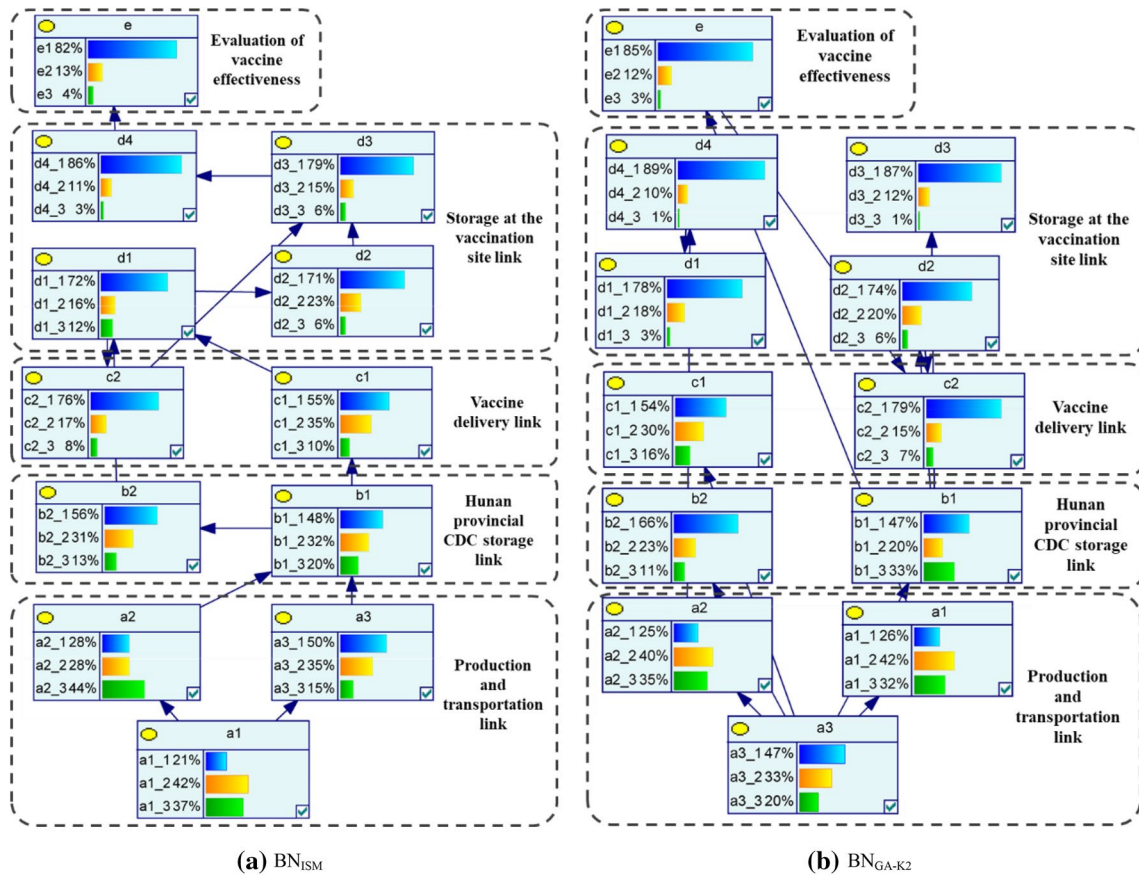
(a) BN_{ISM}

(b) BN_{GA-K2}

Fig. 9 Comparison of Single Structure BN for VE Risks: **a** description of BNISM; **b** description of BNGA-K2

Table 6 Comparison of experimental results between BNISM and BNGA-K2

Accuracy /%	Prediction object				
	c1	d2	d3	d4	c2
BN$_{ISM}$	73.4	75.8	82.8	65.3	68.3
BN$_{GA-K2}$	78.4	80.3	87.4	86.2	87.1

The data was sampled eight times, with 200 sets of random data taken as a group each time. In addition to the neural network models of the original manuscript, two more methods, decision tree methods and naive Bayesian network are slected for the comparison with the proposed Bayesian networks. Figure 10 in the revised manuscript illustrates the accuracy of the results obtained from these four methods. The horizontal axis of the graph indicates the number of trials and the vertical axis shows the difference in accuracy from 100%. The neural network is occasionally able to produce very accurate results, but has the disadvantage of unstable and wabbly results, which need to be optimised. The accuracy of the naive Bayesian network does not compete
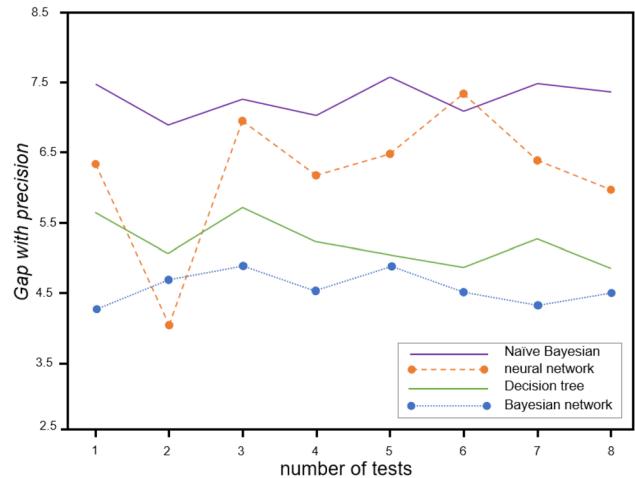

Fig. 10 Accuracy gap comparison of VE prediction

with the results obtained from the Bayesian network method due to the improvement in this paper. Moreover, the Bayesian network is more interpretable. We compared the experimental results of the proposed Bayesian Network method with the decision tree model, which is also interpretable, and

found that the Bayesian network is more accurate, averaging 0.5% than the decision tree model.

3.3 Probabilistic reasoning

Probabilistic reasoning is to use the values of some nodes in the BN (called evidence) to infer the posterior distribution of unknown nodes (called observation nodes). The BN has a bidirectional reasoning function: Reason about the probability of the resulting child nodes from the parent nodes as the cause (prediction reasoning), and infer the probability of parent nodes by using the information of their child nodes (diagnostic reasoning).

Perform analysis of the control variables in the BN model (Fig. 7). First, the predictive reasoning is performed: Assume that the values of the three variables in the production and transportation link are all 1, that is, the GDP ranking of the vaccine manufacturers' cities last year accounted for the top 30% of the province's GDP rankings, the comprehensive quality of refrigerated truck drivers is good, and the temperature of refrigerated truck compartments is fully qualified. The effect of changes in VE values is small. Conversely, if all the three variables of the factory transportation link take the value 3, that is, all of them fail, VE will decrease by 5%. Assume that the values of the four variables in the Hunan provincial CDC storage and vaccine delivery link are 3, that is, the storage room temperature exceeds 8℃ for more than half an hour, the comprehensive quality of storage guards is lower, personal qualities of delivery staff of vaccine delivery link is lower, and the refrigerator compartment temperature of vaccine delivery link exceeds 8 ℃, for more than half an hour, VE will reduce by 6%. Assume that the values of all factors in the storage at the vaccination site are all 3, that is, the comprehensive quality of the vaccination nurses is lower, the temperature in the storage room of vaccination sites exceeds 8 ℃ for more than half an hour, the city where the vaccination site is located ranking the last 40% of the province's GDP in the last year, and the highest temperature on the day of inoculation is above 30 ℃, VE will decrease by 10%. This shows that the last three links have a greater impact on VE than the first link, and it is recommended that the regulatory authorities pay more attention to strengthening the supervision of these three links.

Then we can perform diagnostic reasoning. Assuming that the antibody titer of the vaccine is zero after one month of vaccination, that is, the test result is unsatisfactory, it then can be inferred that the probability of the highest temperature above 30 ℃ on the day of vaccination is 73%, the probability that the temperature of the refrigerated truck compartment during the distribution process exceeds 8 ℃ for more than half an hour is 19%, the probability of a lower overall quality of refrigerated truck drivers in the production

and transportation was 44%, and the probability of lower comprehensive quality of storage guards is 13%. Studies have shown that in addition to the storage temperature of the vaccine, the daily temperature of the vaccine injection, the economic level of the vaccine production site, and the quality of staff handling vaccines all have varying degrees of impact on VE Therefore, the investigation can be performed according to the order of probability to quickly determine the risk factors for vaccine failure, to carry out fast and accurate accountability.

Surprisingly, when performing the single-variable control, no matter what value of node a1 of the GDP ranking of the vaccine manufacturers' cities is set, the probability of the VE node does not change, indicating that there is no direct correlation between the GDP ranking of the vaccine manufacturers 'cities and the VE If the value of the temperature on the day of vaccination is set to 3, which means the maximum daily temperature is above 30 ℃, the probability of VE will decrease by 3%. If the value of the comprehensive quality of staff handling vaccines (a2, b2, c1, d1) is all set to 1, the probability of VE will increase by 6%, which indicates that the comprehensive quality of staff handling vaccines has a significant impact on VE

Different from the existing studies that only focused on the temperature of the cold chain of vaccines (Pourhoseingholi et al. 2012; Das et al. 2019), this study incorporates the economic level of vaccination sites' cities, the temperature on the day of injection, the comprehensive quality of staff handling vaccines, and scores the total quality quantification into one overall measure, which has not been proposed in previous studies. Although our experiment is a randomized experiment, which theoretically eliminates the influence of confounding factors, due to the consideration of experimental rigor, there may be other non-measurable factors. However, from the results of our experiments, BN has excellent interpretability and accuracy in the analysis of VE and inference.

Further, if the BN model is used for diagnostic reasoning, it can quickly and effectively track the adverse events of ineffective vaccines, which will help improve VE and reduce the spread of infectious diseases.

4 Conclusion

This paper designs a BN model that requires only limited historical data and expert questionnaires to perform more accurate VE analysis and empirical analysis. In this study, we focused on the influenza vaccine in the second-class vaccine. Combined with historical literature and domain expert experience, and because of the impact of vaccine storage temperature on VE, eleven possible risk factors, existing in

the four main stages of the cold chain transportation links, were analyzed.

In cooperation with the Hunan Provincial CDC, through data collection and data discretization, the ISM-K2 BN model was adopted to conduct a risk assessment of VE Findings help to clarify the cold chain distribution responsibility of vaccines and accurately strengthen the requirements of each link in "Transportation-Storage-Distribution-Inoculation" separately.

The study can provide individual assistance for the guarantee and evaluation of VE, and serve as a specific reference for strengthening China's vaccine quality and safety supervision value. The BN model, combined with the ISM method, provides a practical reasoning approach for analyzing vaccine-related data obtained from the supply chain process, and in practice, serves as a prediction strategy for vaccine effectiveness risk.

As for the limitation of the study, for one thing, some factors that affect VE, such as unknown manufacturers' qualifications, the individual physique of the vaccine recipients, were not included in the models. For another, information is nonadditive, and due to the rarity of events, identification of point probabilities or probability distributions to characterize parameter uncertainty would be subject to degrees of imprecision. In such cases, we will express the assessment of an event in the form of interval or imprecise probabilities using Dempster-Shafer's belief function and Choquet capacity. These aspects need to be further improved in future research.

Declarations

Conflict of interest All authors declared that they have no conflict of interests.

Informed consent Any form of biomedical, clinical and biometric data included in the manuscript has obtained the subject's informed consent.

Ethical approval The study was conducted in accordance with the guidelines of relevant laws and institutions and approved by the corresponding ethics committee.

References

Boutselis P, McNaught K (2019) Using Bayesian networks to forecast spares demand from equipment failures in a changing service logistics context. Int J Prod Econ 209:325–333

Brison M, LeTallec Y (2017) Transforming cold chain performance and management in lower-income countries. Vaccine 35:2107–2109

CDC (1999) Ten great public health achievements—United States, 1900–1999. In: Mmwr Morbidity & Mortality Weekly Report

Cooper GF, Herskovits E (1992) A Bayesian method for the induction of probabilistic networks from data. Mach Learn 9:309–347

Das MK, Arora NK, Mathew T, Vyas B, Sindhu M, Yadav A (2019) 'Temperature integrity and exposure to freezing temperature during vaccine transfer under the universal immunization program in Three States of India. Indian J Public Health 63:139

Dasaklis TK, Pappis CP, Rachaniotis NP (2012) Epidemics control and logistics operations: a review. Int J Prod Econ 139:393–410

Falchi A, Souty C, Grisoni ML, Mosnier A, Hanslik T, Daviaud I, Varesi L, Kerneis S, Carrat F, Blanchon T (2013) Field seasonal influenza vaccine effectiveness: evaluation of the screening method using different sources of data during the 2010/2011 French influenza season. Hum Vaccine Immunother 9:2453–2459

Foppa IM, Haber M, Ferdinands JM, Shay DK (2013) The case test-negative design for studies of the effectiveness of influenza vaccine. Vaccine 31:3104–3109

Ganz PA, Shanley JD, Cherry JD (1978) Responses of patients with neoplastic diseases to influenza virus vaccine. Cancer 42:2244–2247

Hanson CM, George AM, Sawadogo A, Schreiber B (2017) Is freezing in the vaccine cold chain an ongoing issue? A literature review. Vaccine 35:2127–2133

Head JR, Collender PA, Lewnard JA, Skaff NK, Li L, Cheng Q, Baker JM, Li C, Chen D, Ohringer A, Liang S, Yang C, Hubbard A, Lopman B, Remais JV (2019) Early evidence of inactivated enterovirus 71 vaccine impact against hand, foot, and mouth disease in a major center of ongoing transmission in china, 2011–2018: a longitudinal surveillance study. Clin Infect Dis 71:3088–3095

Heckerman D, Geiger D, Chickering DM (1995) Learning Bayesian networks: the combination of knowledge and statistical data. Mach Learn 20:197–243

Hosseini S, Barker K (2016) A Bayesian network model for resilience-based supplier selection. Int J Prod Econ 180:68–87

Ivanov D, Dolgui A (2020) OR-methods for coping with the ripple effect in supply chains during COVID-19 pandemic: managerial insights and research implications. Int J Prod Econ 107921

Jackson ML, Nelson JC (2013) The test-negative design for estimating influenza vaccine effectiveness. Vaccine 31:2165–2168

Kinfe Y, Gebre H, Bekele A (2019) 'Factors associated with full immunization of children 12–23 months of age in Ethiopia: a multilevel analysis using 2016 Ethiopia Demographic and Health Survey. PLoS ONE 14:e0225639

Kumru OS, Joshi SB, Smith DE, Russell Middaugh C, Prusik T, Volkin DB (2014) Vaccine instability in the cold chain: mechanisms, analysis and formulation strategies. Biologicals 42:237–259

Pearl J, Robins JM, Greenland S (1999) Confounding and collapsibility in causal inference. Stat Sci 14:29–46

Phillips DE, Dieleman JL, Lim SS, Shearer J (2017) Determinants of effective vaccine coverage in low and middle-income countries: a systematic review and interpretive synthesis. BMC Health Serv Res 17:681

Pourhoseingholi MA, Baghestani AR, Vahedi M (2012) How to control confounding effects by statistical analysis. Gastroenterol Hepatol Bed to Bench 5:79–83

Rogers B, Dennison K, Adepoju N, Dowd S, Uedoi K (2010) "Vaccine cold chain: part 2 training personnel and program management." AAOHN J 58:391–400

Rupprecht CE, Salahuddin N (2019) Current status of human rabies prevention: remaining barriers to global biologics accessibility and disease elimination. Expert Rev Vaccines 18:629–640

Strohfus PK, Kim SC, Palma S, Duke RA, Remington R, Roberts C (2017) Immunizations challenge healthcare personnel and affects immunization rates. Appl Nurs Res 33:131–137

Torvaldsen S, Mcintyre PB (2002) Observational methods in epidemiologic assessment of vaccine effectiveness. Commun Dis Intell Q Rep 26:451–457

Tostanoski LH, Jewell CM (2017) Engineering self-assembled materials to study and direct immune function. Adv Drug Del Rev S0169409X17300352

Vaillant L, La Ruche G, Tarantola A, Barboza P (2009) Epidemiology of fatal cases associated with pandemic H1N1 influenza 2009. J Euro Survll 14

Ventola CL (2016) Immunization in the United States: recommendations, barriers, and measures to improve compliance: part 1: childhood vaccinations. P & T A Peer Rev J Formul Manag 41:426–436

Warfield JN (2010) Developing Interconnection matrices in structural modeling. Systems ManCybernetics IEEE Transactions on, SMC-4, pp 74–80

Weinberg GA (2010) Vaccine epidemiology: efficacy, effectiveness, and the translational research roadmap. J Infect Dis 201:1607–1610

WHO (2004) Immunization in practice. Kyushu Sangyo Univ Bus Rev 11:45–57

WHO (2020a) 10 facts on immunization. WHO (World Health Organization). Accessed May. 5. https://www.who.int/features/factfiles/immunization/en/

WHO (2020b) Coronavirus disease (COVID-2019) situation reports. https://www.who.int. Accessed 5 May

WHO (2020c) Ebola virus disease—democratic Republic of the Congo. https://www.who.int/csr/don/13-february-2020c-ebola-drc/en/. Accessed 5 May

WHO (2020d) From warehouse to remote indigenous communities: The journey of vaccines in Brazil. https://www.who.int/news-room/feature-stories/detail/from-warehouse-to-remote-indigenous-communities-the-journey-of-vaccines-in-brazil. Accessed 5 May

WHO (2020e) Middle East respiratory syndrome coronavirus (MERS-CoV) – United Arab Emirates. https://www.who.int/csr/don/31-january-2020e-mers-united-arab-emirates/en/. Accessed 5 May

WHO (2020f) Summary of probable SARS cases with onset of illness. https://www.who.int/csr/sars/country/table2004_04_21/en/. Accessed 5 May

WHO (2020g) WHO vaccine management handbook module VMH-E2–01.1. How to monitor temperatures in the vaccine supply chain. http://apps.who.int/iris/bitstream/10665/183583/1/WHO_IVB_15.04_eng.pdf. Accessed 5 May

Yakum MN, Ateudjieu J, Pélagie FR, Walter EA, Watcho P (2015) Factors associated with the exposure of vaccines to adverse temperature conditions: the case of North West region, Cameroon. BMC Res Notes 8:277

Yang S, Tian G, Cui Y, Ding C, Deng M, Chengbo Yu, Kaijin Xu, Ren J, Yao J, Li Y, Cao Q, Chen P, Xie T, Wang C, Wang B, Mao C, Ruan B, Tian'an Jiang, and Lanjuan Li. (2016) Factors influencing immunologic response to hepatitis B vaccine in adults. Sci Rep 6:27251

Zapata H (2017) Multicohort analysis reveals baseline transcriptional predictors of influenza vaccination responses HIPC-CHI Signatures Project Team* and HIPC-I Consortium. Sci Immunol

Zhou Y, Fenton N, Zhu C (2016) An empirical study of Bayesian network parameter learning with monotonic influence constraints. Decis Support Syst 87:69–79

Zhu Qi, Hao YuanTao, Ma JiaQi, ShiCheng Yu, Wang Yu (2011) Surveillance of hand, foot, and mouth disease in Mainland China (2008–2009). Biomed Environ Sci 24:349–356

Zhuang W-H, Wang Y-P (2020) Analysis of the immunity effects after enhanced hepatitis B vaccination on patients with lymphoma. Leuk LymphoMa 61:357–363