



OPEN

Construction and effectiveness evaluation of a knowledge-based infectious disease monitoring and decision support system

Mengying Wang¹, Mo Jia¹, Zhenhao Wei², Wei Wang², Yafei Shang² & Hong Ji¹✉

To improve the hospital's ability to proactively detect infectious diseases, a knowledge-based infectious disease monitoring and decision support system was established based on real medical records and knowledge rules. The effectiveness of the system was evaluated using interrupted time series analysis. In the system, a monitoring and alert rule library for infectious diseases was generated by combining infectious disease diagnosis guidelines with literature and a real medical record knowledge map. The system was integrated with the electronic medical record system, and doctors were provided with various types of real-time warning prompts when writing medical records. The effectiveness of the system's alerts was analyzed from the perspectives of false positive rates, rule accuracy, alert effectiveness, and missed case rates using interrupted time series analysis. Over a period of 12 months, the system analyzed 4,497,091 medical records, triggering a total of 12,027 monitoring alerts. Of these, 98.43% were clinically effective, while 1.56% were invalid alerts, mainly owing to the relatively rough rules generated by the guidelines leading to several false alarms. In addition, the effectiveness of the system's alerts, distribution of diagnosis times, and reporting efficiency of doctors were analyzed. 89.26% of infectious disease cases could be confirmed and reported by doctors within 5 min of receiving the alert, and 77.6% of doctors could complete the filling of 33 items of information within 2 min, which is a reduction in time compared to the past. The timely reminders from the system reduced the rate of missed cases by doctors; the analysis using interrupted time series method showed an average reduction of 4.4037% in the missed-case rate. This study proposed a knowledge-based infectious disease decision support system based on real medical records and knowledge rules, and its effectiveness was verified. The system improved the management of infectious diseases, increased the reliability of decision-making, and reduced the rate of underreporting.

Infectious diseases have always been a critical public health challenge for human society. The COVID-19 pandemic posed a substantial threat to the life and health of all mankind and caused considerable economic losses to countries around the world^{1,2}. As the first line of defense in the diagnosis and treatment of infectious diseases, hospitals' real-time monitoring and early warning capabilities for infectious diseases have a major impact on the success and efficiency of the entire infectious disease prevention and control work. Early infectious disease warning and reminders can affect the treatment effect and prevention and control of infectious diseases considerably^{3,4}. Studies^{5,6} have mainly relied on laboratory test results for monitoring and early warning. For example, during the COVID-19 pandemic, medical professionals and hospital managers were alerted by critical values when nucleic acid test results were positive⁷. Previous studies^{8,9} have shown that electronic reminders based on test results have important implications for disease diagnosis and treatment, especially in the field of infectious diseases, as they enable early isolation and management of infectious diseases. However, relying solely on test results for accurate and comprehensive infectious disease early warning is inadequate, as infectious disease diagnosis requires information from multiple sources, such as symptoms, signs, diagnosis, and medical history throughout the patient's entire medical record. With the gradual proliferation of Electronic Health Record Systems (EHRS), EHRS has become a management system for daily applications in hospital outpatient and inpatient departments. Several studies have shown that EHRS can accelerate clinical information flow, promote medical data integration, and

¹Information Management and Big Data Center, Peking University Third Hospital, Beijing, China. ²Goodwill Hessian Health Technology Co. Ltd, Beijing, China. ✉email: puh3_imc@bjmu.edu.cn

improve the efficiency and quality of medical services^{10,11}. However, current EHRS are mainly designed for clinical practice rather than public health purposes¹². Therefore, it is difficult to directly monitor infectious diseases according to relevant diagnostic criteria using EHRS, and manual retrieval and report generation are required, which is time-consuming and prone to omissions.

After the outbreak of SARS in 2003, in addition to monitoring infectious diseases, China established a direct reporting system for infectious diseases. China divides infectious diseases into three categories, A, B, and C, based on their impact and harmfulness, with A being the highest priority. When a case is diagnosed as a certain infectious disease, it is required to be reported to the Chinese Infectious Diseases Reporting System (CIDRS) within a designated time frame. Clinical doctors are required to report to the hospital's disease control management department through an electronic medical record system, after which the hospital's disease control department reviews the reported information before transcribing it into CIDRS¹³. Unfortunately, given the lack of uniformity among the electronic medical record systems of various hospitals, it is difficult to connect them all to the national system, and transcription reports involve time-consuming and error-prone manual operations. At present, for infectious disease reporting based on electronic medical records, some medical institutions in China choose to judge whether a disease is an infectious disease by monitoring diagnosis names, and monitor cases that are diagnosed as infectious diseases but are not actually reported to a nationwide authority.

In order to make electronic medical record systems more intelligent, preliminary research¹⁴ extended Clinical Decision Support Systems (CDSSs) on the basis of electronic medical records. A CDSS is a system that provides intelligent processing of disease-specific and patient-specific data for clinical physicians, patients, and other individuals¹⁵. Several diverse types of CDSSs can be found in literature^{16,17}, such as diagnostic tools, expert systems, and workflow support. CDSSs have provided considerable assistance in clinical practice, with over 62% of trials demonstrating enhancements in practitioner performance, reminder systems, drug dosage systems, and disease management systems. At present, more than one-third of CDSS applications are mainly used in chronic disease management, treatment prediction, and risk assessment¹⁷. For instance, Dalal et al.¹⁸ and Khan et al.¹⁹ developed treatment risk alerts for elderly patients, considering factors such as suitability for catheter insertion, likelihood of physical function recovery, and potential drug interactions. In Tao et al.'s study²⁰, CDSS was shown to reduce the average length of hospital stays by almost one day, demonstrating their potential to effectively enhance medical efficiency and improve clinical practice. In contrast, some studies have analyzed how CDSS introduction interferes with physician workflows, with ineffective alarms and frequent alerts leading to physician fatigue and oversight²¹.

With respect to the alerting methods used in CDSS applications, Park et al.'s research²² used interruptive alerts in drug interaction warnings, which directly interrupted the clinical physician's workflow with pop-up notifications. Binh et al.²³ used list-style alerts to remind clinical physicians to complete preventive care orders. The alerting methods in previous research were all single-mode reminders, lacking the ability to set varying levels of alerts. As for the interpretability of CDSS, in our preliminary research²⁴, we employed the Multiple Infectious Disease Diagnostic Model (MIDDM) to provide warning alerts based on medical records. However, owing to the limitations of machine learning in providing evidence-based explanations, clinical physicians lack reliable bases for alerts, which can have noticeable impacts on patients, especially in cases of misdiagnosing infectious diseases.

As of now, CDSS applications with better interpretability are based on rule-based knowledge bases. However, previously, CDSS rule libraries were established based on clinical experts, which were then hardcoded by technical developers into the system's programming backend, making them difficult to modify. Thus, CDSS rules are challenging to develop, update, and maintain²⁵. Consequently, one of the concerns to be addressed in this research is how to extract rules based on clinical knowledge as well as how to supplement, update, and maintain these rules through the system.

Therefore, in order to address the problems of passive and single-diagnosis or single-test infectious disease monitoring, and the complete dependence on manual active reporting of infectious diseases, this study establishes an infectious disease monitoring and decision support system to assist clinical physicians in infectious disease decision-making. Based on clinical guidelines and real medical record knowledge maps, complex rules are generated to simultaneously monitor patient medical records, examination reports, and other results, and comprehensively judge whether they meet the national diagnostic criteria for infectious diseases. This meets the needs of infectious disease monitoring and alerts doctors to cases that meet the positive indicators of infectious diseases, assisting them in making diagnostic decisions. In addition to physician alerts, a management backend is also established for management departments to manage and monitor infectious disease reporting.

Methods

System design and implementation. This study focused on a large comprehensive hospital and its three branch hospitals in China, having a total of 2353 beds and a daily average outpatient and inpatient volume of 17,000. Based on the hospital's electronic medical record system and laboratory information system, an Infectious Disease Monitoring and Decision Support System was established. The system monitors all patient medical record information during their visit, including age, gender, vital signs, medical history, laboratory test results, diagnosis, and radiographic examination results. As the hospital system generates a random ID for each patient and saves it in a secure server area, the anonymity of all patients was maintained. Data analysis was conducted using Excel to calculate the proportion of time intervals. The R 4.2.2 software was used to analyze the changes in slope before and after the system's implementation as well as the immediate level changes at the intervention points. The computers used by hospital doctors had the same configuration, i.e., Intel(R) Core(TM) i7-7500U CPU @ 2.70 GHz, 8 GB of running memory, and a total hard disk size of 256 GB.

The study was approved by the Medical Science Research Ethics Committee of Peking University Third Hospital (serial number: IRB00006761-M2022287). The Peking University Third Hospital Medical Science Research

Ethics Committee approved our study and waived the need for informed consent. We confirm that all the experiment protocols for including humans were in accordance with the guidelines of national/international/institutional requirements or the Declaration of Helsinki.

System architecture and workflow design. The hospital selected for this study has a good level of informatization and has aggregated data from multiple systems, such as EMR, medical orders, laboratory tests, and inspections, to the data center. All unstructured data are processed into structured data in the data center. The Infectious Disease Monitoring and Decision Support System can directly connect to the data center to obtain the full range of data, including medical records and laboratory tests. The system consists of two parts: the decision support end (AI assistant) for doctors and the monitoring and management end for the hospital's disease control department. The doctor's decision support end (AI assistant) is integrated with the outpatient and inpatient doctor station, and the AI assistant system is automatically opened when the doctor logs in to the doctor station, with the listening status indicated by a small AI Assistant icon. The monitoring and management end of the hospital's disease control department is a web-based page that can only be accessed by authorized personnel and is dedicated for management department use. The system data flow is shown in Fig. 1.

Establishment of a complex rule library based on knowledge map and diagnostic criteria. Infectious disease diagnoses can be classified into suspected cases, clinically diagnosed cases, confirmed cases, and carriers of pathogens^{26,27}. First, based on diagnostic criteria, two groups of senior experts from infectious disease departments such as respiratory, pediatrics, and dermatology summarized the guidelines and then cross-verified and discussed to form standardized content corresponding to the medical record data source. For example, the rule extraction for influenza is shown in Fig. 2. In Fig. 2, the extracted rule criteria are based on "Standards for Diagnosis of Statutory and Other Key Monitored Infectious Diseases in the Health Industry of the People's Republic of China". The expert group conducting cross-validation found it challenging to list all the symptoms or physical examination descriptions of doctors in one go during symptom supplementation. This led to several omissions in the monitoring process. Therefore, considering the actual operation of the hospital, this study uses a graph generated based on real historical data to analyze the doctor's writing habits and actual medical record data for rule revision and supplementation, so that the monitoring rules can be triggered in real clinical work.

In this study, we employed natural language processing techniques²⁸ established in previous research as the foundation. For the non-structured portion of real-world clinical data, we utilized the BiLSTM-CRF network²⁹, an open-source method with effective sequence labeling performance. We combined this with rule-based models and other methods to extract information from the original electronic medical records. In this process, first, tokenized and serialized text were input into the BiLSTM layer, combining the results of forward and backward hidden states to generate the BiLSTM output³⁰. Subsequently, the output of BiLSTM was fed into the CRF layer, forming the BiLSTM-CRF network structure. This architecture takes advantage of both BiLSTM and CRF: the BiLSTM, as a bidirectional LSTM component, effectively preserves the contextual information of the entire sentence, enabling feature extraction from the sentence, while the CRF layer leverages the ability to learn constraint information from the training data, further enhancing the accuracy of information extraction, as illustrated in Fig. 3.

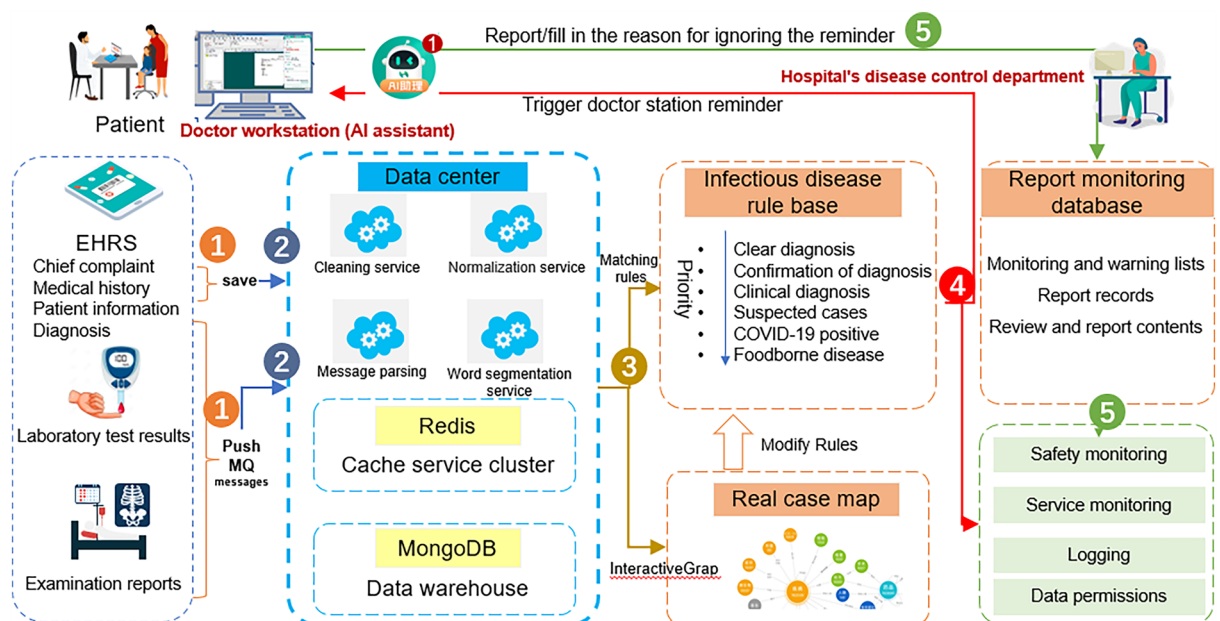


Figure 1. Data flow of the infectious disease monitoring and decision support system.

Original diagnostic criteria		Summary of the diagnostic criteria corresponding to the early warning system	Trigger conditions
clinical diagnosis A history of epidemiological exposure (within 7 days before onset, close contact with suspected or confirmed influenza patients without effective personal protection, or belonging to one of the clustered cases of influenza-like cases, or there is clear evidence of infection transmission to others) and the above clinical manifestations of influenza, with exclusion of other diseases that may cause influenza-like symptoms.		<ol style="list-style-type: none"> Chief complaint includes fever. Current medical history - respiratory symptoms including nasal congestion, cough, runny nose, or sore throat. Current medical history - systemic symptoms including fatigue or muscle soreness. Department name includes or is related to Respiratory Department or Inpatient - Respiratory Department, Outpatient - Fever Emergency, Outpatient - Respiratory Department, Emergency Department - Internal Medicine, Outpatient - Infectious Diseases Department. 	(1) Suspected diagnosis 1, 2, 3, 4. (2) Confirmed diagnosis with positive results on laboratory test items 5, 6, 7, or 8.
Clinical manifestations Mainly manifested as fever, headache, myalgia, and malaise, with a body temperature of up to 39-40°C, accompanied by chills, shivering, and other systemic symptoms such as muscle and joint pain, fatigue, and decreased appetite. There may also be sore throat, dry cough, nasal congestion, runny nose, discomfort behind the sternum, facial flushing, conjunctival congestion, etc. Some patients may have mild or no flu-like symptoms. Children infected with type B influenza often have vomiting, abdominal pain, and diarrhea as the main symptoms.			
Confirmed diagnosis Having the above clinical manifestations of influenza, with one or more of the following positive results in pathogen detection: 1. Positive influenza virus nucleic acid detection; 2. Positive influenza antigen detection; 3. Positive influenza virus culture; 4. A four-fold or higher increase in the level of influenza virus-specific IgG antibodies in paired acute and convalescent serum.			
		<ol style="list-style-type: none"> Laboratory test items include influenza A virus nucleic acid detection, with qualitative results including positive or +. Laboratory test items include influenza A virus antigen detection, with qualitative results including positive or +. Laboratory test items include influenza B virus antigen detection, with qualitative results including positive or +. Laboratory test items include influenza B virus ribonucleic acid (RNA) detection, with qualitative results including positive or +. 	

Configure rules in the system

Number	ID	规则内容 表达式 元素 Rule contents
13013	448855778833228 80066FF2266CC00 DD	(Current department name) contains Respiratory Internal Medicine or Pediatrics; (Symptom name) contains fatigue or muscle soreness or muscle pain or fever. (Name of Laboratory Test Item.) Equals to: "Influenza A virus nucleic acid test" OR "Influenza A virus antigen test" OR "Influenza B virus antigen test" OR "Influenza B virus test" OR "Influenza A virus H1N1 subtype detection" OR "Influenza B virus" OR "Influenza B virus nucleic acid" (Qualitative result of the laboratory test) "+" or positive.

Figure 2. Records of rules extracted for influenza and visualization of rule configuration in the system.

Based on the structured data, a knowledge map was drawn using the open-source InteractiveGraph graph data interactive visualization analysis framework³¹, as shown in Fig. 4. Separate graphs are drawn for outpatient and inpatient cases, displaying the symptoms and doctors' notes that appear in the real medical records pertaining to an infectious disease. For high-frequency symptoms, larger graphics are used, and the occurrence frequency is indicated on the connecting line. Considering the readability of the graph and the rotation of doctors, the data is displayed for the past six months, making it convenient for rule review experts to supplement rules at any time.

Decision support for various types of infectious diseases. The system provides infectious disease warning reminders of various severity levels based on the rule library. The AI assistant on the doctor's decision-making end is integrated with the doctor's workstation. When a doctor saves a medical record, the text structuring service is called immediately to perform word segmentation processing. It then matches a variety of level rules in the knowledge rule library and makes a diagnosis directly (the doctor enters the diagnosis name directly in the medical record diagnosis). The highest priority is given to a direct diagnosis, followed by confirmed diagnosis, clinical diagnosis, and suspected cases. The rule matching and decision support process is shown in Fig. 5, and the number of rules is shown in Fig. 6. The number of rules varies considerably for different diseases owing to the influence of their clinical diagnostic guidelines and disease complexity. Diseases such as epidemic encephalitis B, measles, and epidemic hemorrhagic fever are associated with a higher number of rules, while diseases like seasonal flu and hand, foot, and mouth disease have fewer rules, each with 27 rules.

Among the diseases, epidemic encephalitis B has the highest number of rules. Specifically, the diagnostic criteria for epidemic encephalitis B are divided into mild, moderate, severe, and extremely severe cases. The rules also involve a combination of physical examination findings and chief complaints or medical history, leading to a considerable increase in the number of rules.

Furthermore, during the COVID-19 pandemic, a new classification was added for monitoring positive COVID-19 test results. This classification has the lowest priority. When new evidence is made available to the clinical diagnosis or confirmed diagnosis criteria, the positive test result will be updated directly under the confirmed diagnosis classification. For example, when the COVID-19 antigen test is positive, it triggers the

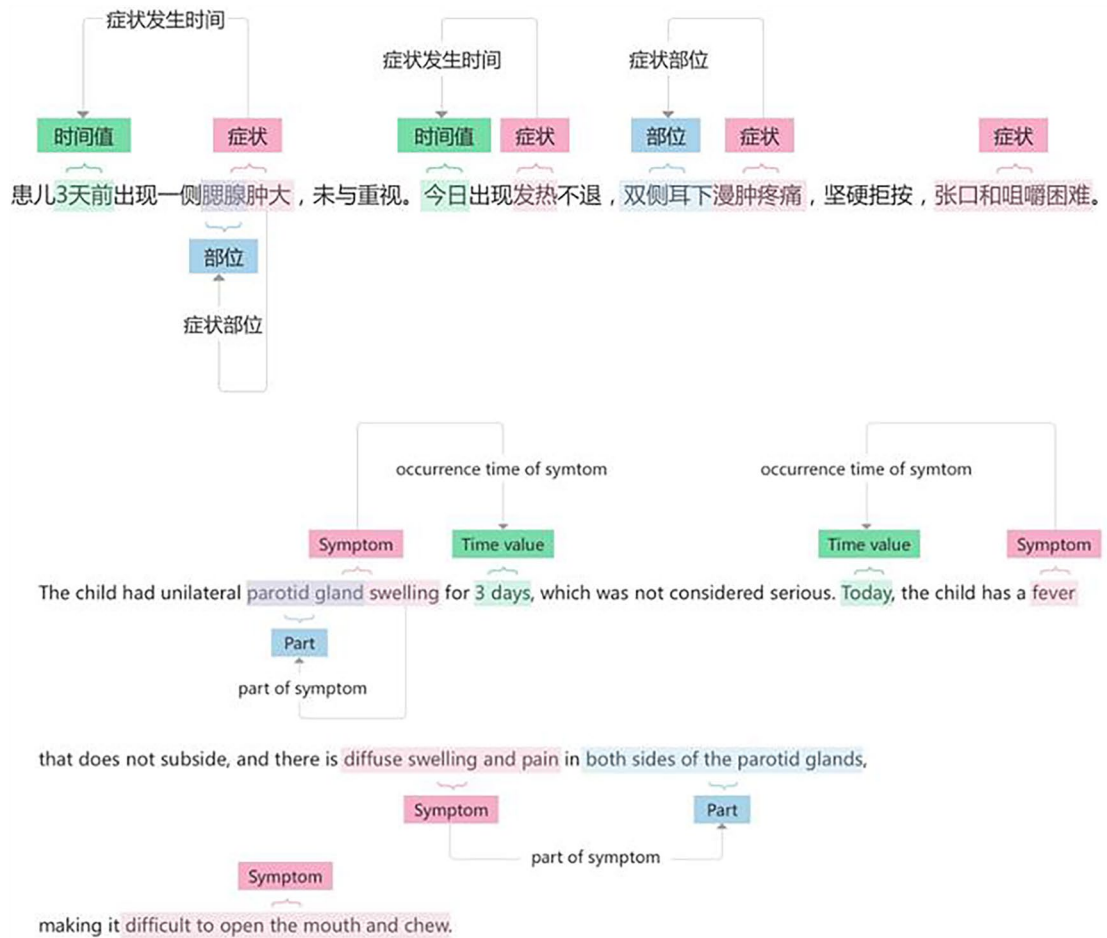


Figure 3. Natural language processing and structured processing of infectious disease medical records (in Chinese and English).

positive monitoring. When the nucleic acid test is positive and the patient has flu-like symptoms, it satisfies the confirmed diagnosis rule, triggering a reminder under the confirmed diagnosis classification, which covers the original positive monitoring.

Clear diagnosis decision support: Doctors issue a clear diagnosis in the doctor’s workstation, and the AI assistant automatically retrieves the diagnostic message. After matching the diagnostic ICD-10 code with the diagnosis set by the infectious disease monitoring rules, the AI assistant will remind the doctor to report the diagnosis once the matching is successful.

Confirmation of diagnosis or clinical diagnosis decision support: Matching the patient’s complaints, medical history and laboratory results, the system provides reminders. For example, for confirmed diagnosis of influenza, when the patient’s test result is positive for influenza virus antigen or “+”, the AI assistant will actively remind the doctor to confirm the patient’s diagnosis of influenza and to diagnose and report the medical record in a timely manner.

Decision support for suspected cases: Suspected cases mainly include those who meet the requirements of the guidelines, such as those with medical history, contact history, or obvious symptoms. The system triggers the rule for suspected cases accordingly. For example, for suspected cases of epidemic parotitis, when the doctor writes “bilateral parotid swelling for 3 days” in the medical record, the system calls the structured service to match the segmented result with the rule library. The AI assistant will actively remind the doctor that the patient is a suspected case of epidemic parotitis and suggest other indicators of the disease. The AI assistant also provides a knowledge base for clinical doctors in non-infectious disease departments to understand infectious diseases.

After the doctor sees the reminder and confirms the diagnosis for the infectious disease medical record that needs to be reported, the system automatically fills in the report content to reduce the doctor’s input workload. The system also automatically checks the data integrity. After the doctor confirms and submits the report, it is sent to the hospital’s disease control department. The infectious disease control personnel monitor the reporting, audit process, and status in real time. For situations where a diagnosis cannot be made and reporting is not required, the system provides an ignore option for interaction between doctors and infectious disease control management personnel. The ignore function avoids missed reporting and duplicate reporting. The reasons for ignoring an infectious disease diagnosis include old lesions and carriers of pathogens that do not need to be

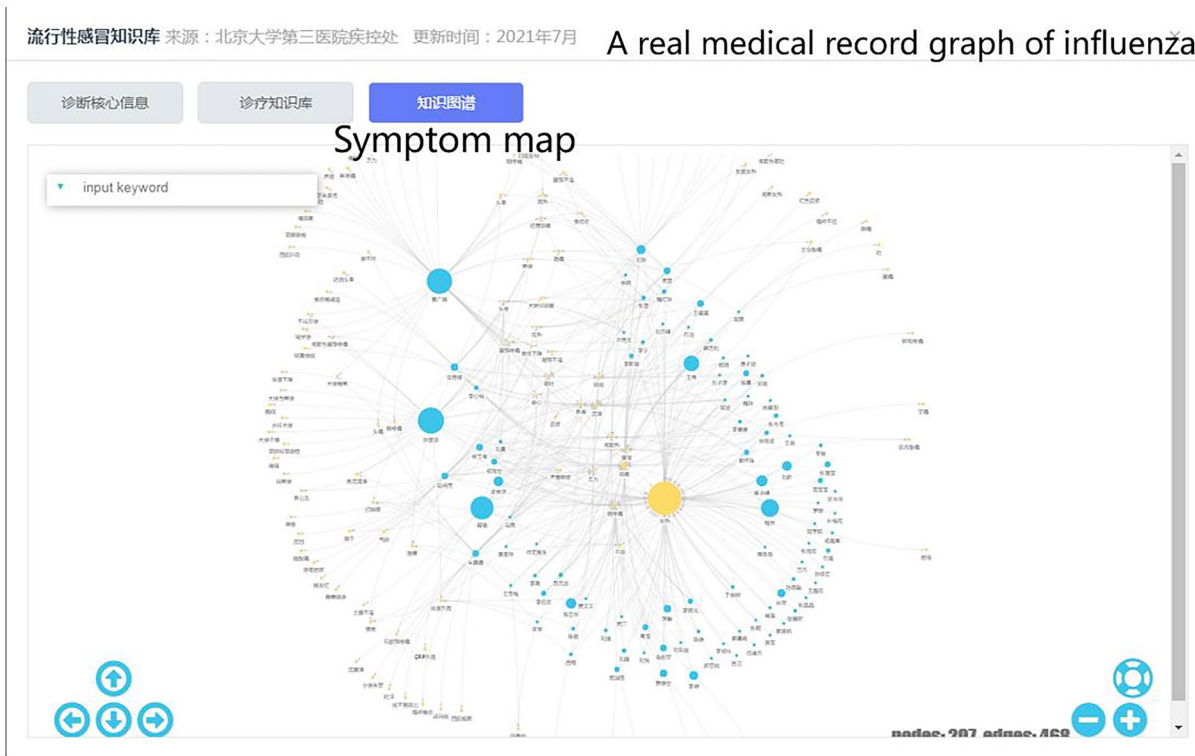


Figure 4. Based on the actual historical medical records of the hospital, a real medical record graph of influenza is generated.

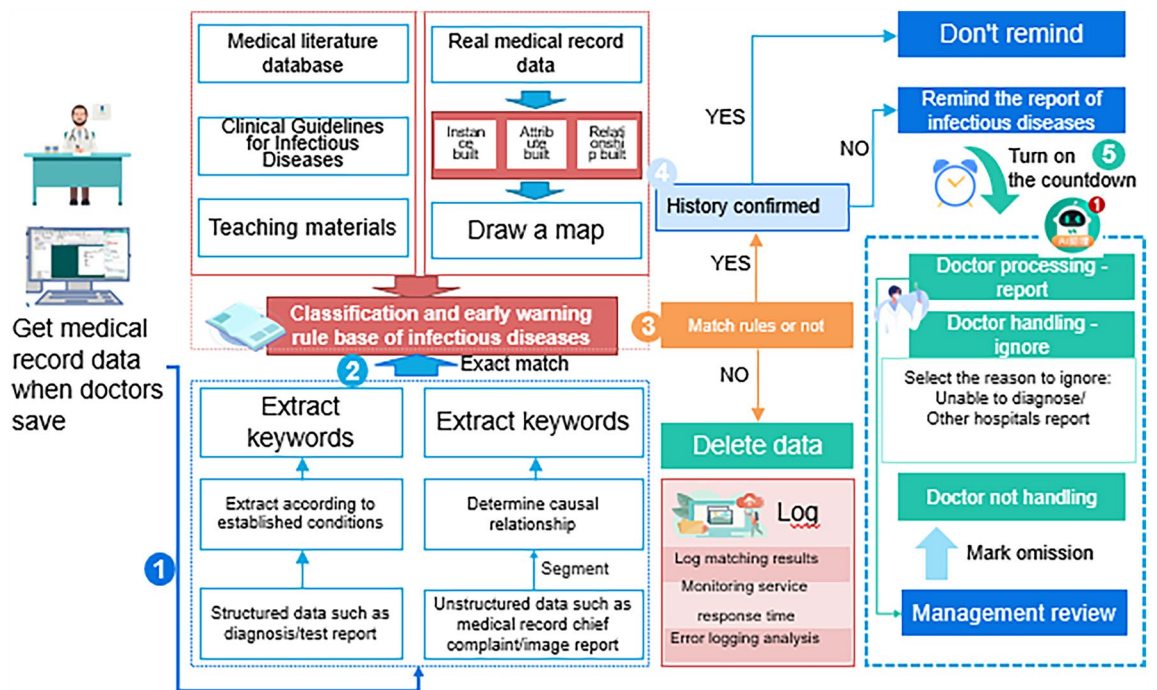


Figure 5. System rule matching process.

reported. This solution effectively solves the problems of missed reporting and misreporting in active reporting methods. The screenshots of the doctor-side and management-side systems are shown in Fig. 7.

Interruption-based and non-interruption-based reminder decision-making. According to China’s classification and management requirements for infectious diseases, medical institutions are required to report Class A infec-

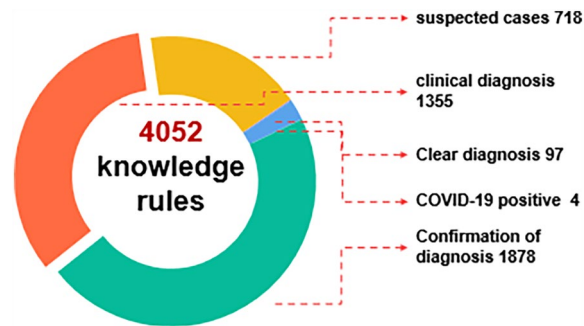


Figure 6. Infectious disease alert rules are counted by type.

tious diseases to CIDRS within 2 h, and Class B and C infectious diseases within 24 h³². Therefore, the reminder system is designed based on the classification and management requirements of infectious diseases. To avoid affecting the doctor's diagnosis of current patients, for rules that require inspection results to trigger, the system triggers infectious disease warning prompts for patients that the doctor has seen when they return to the patient list after ending the current patient's consultation. When a reminder needs to be sent to other doctors for follow-up consultation or reporting, the system automatically combines the reporting and reminder results to avoid duplication. A variety of reminder methods are shown in Fig. 8.

Popup interrupt reminder: For Class A infectious diseases that have already been diagnosed, the decision support system interrupts the doctor's work directly in the form of a popup window, and the small icon of the decision support system expands to show details, highlighting and flashing the reason for the interruption, such as "Monitoring reason: diagnosis includes Pestis, disposition reminder: report required within 2 h", to draw enough attention from the doctor.

Non-interruptive bubble reminder: For Class B and C infectious diseases, after a clear diagnosis has been made, the AI assistant icon of the decision support system appears in a bubble dialogue box, reminding the doctor to report the infectious disease and automatically starting a 24-h countdown. The doctor can click on the AI assistant to expand the report card for reporting. For infectious diseases that meet the clinical diagnosis and suspected conditions mentioned in "[Decision support for various types of infectious diseases](#)" but do not have a clear diagnosis, the bubble dialogue reminder is still triggered. After the doctor clicks on the AI assistant, they can view the specific rule content that is met, such as "Monitoring reason: chief complaint includes bilateral parotid gland enlargement, disposition reminder: report or ignore". The doctor can see this reminder until it is resolved. After the doctor reports or fills in the reason for ignoring, the disease control manager can view the interaction content in the monitoring system.

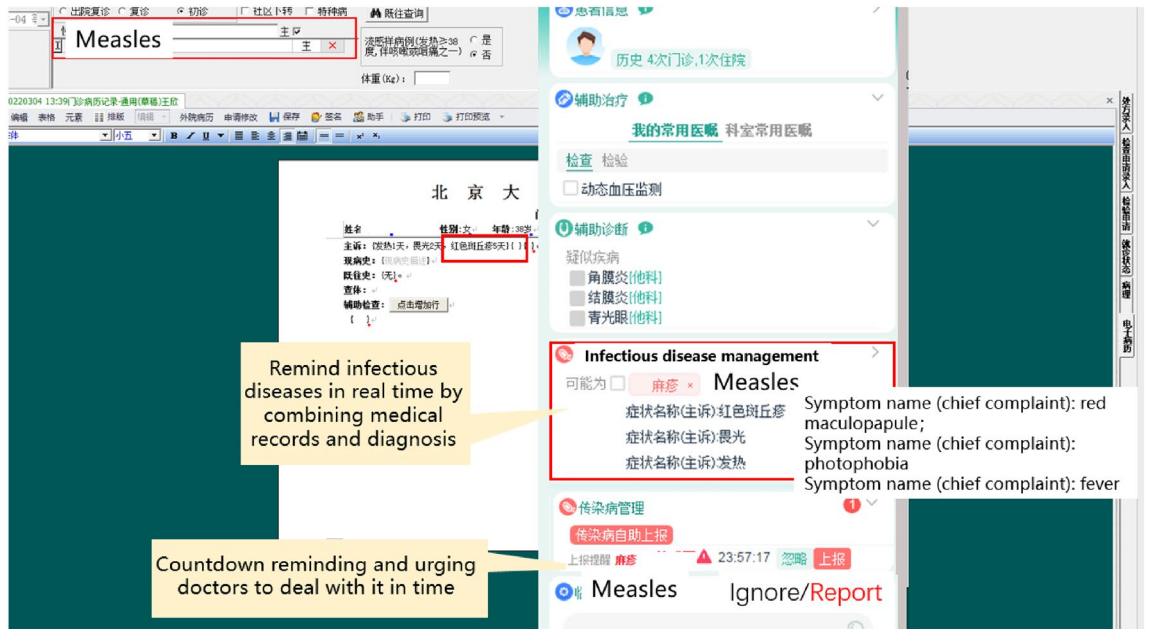
No reminder: For patients who have reported infectious diseases within the last 30 days, no reminder is given during their follow-up consultations.

Management department monitoring and inspection. For the hospital's disease control department, the system mainly meets the requirements of in-hospital infectious disease monitoring, content review, and abnormal symptom monitoring and analysis. For content review, disease control department staff can revise the content and generate a revision mode. Doctors can view the reasons for the modification, which improves the quality of reported content. For infectious disease monitoring, a list of detailed reasons for triggering rules at all doctor stations in the hospital will be displayed. Disease control department staff can mark and make notes on cases of missed reporting where doctors have not processed the reminders and alerts.

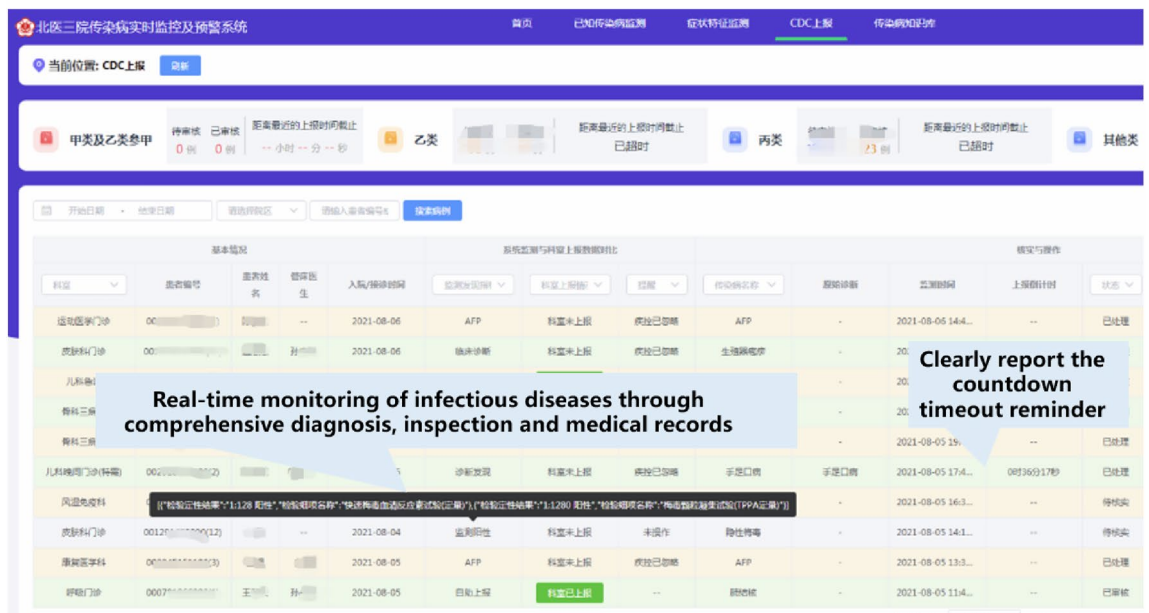
Reducing the intervention of invalid alarms. Setting trigger restrictions for specific diseases and specialties: For sexually transmitted diseases such as chlamydia trachomatis infection and gonorrhea, diagnosis and reporting must be done by specialists with diagnostic privileges in dermatology, obstetrics and gynecology, and urology, as per the diagnostic guidelines. For hand, foot, and mouth disease, diagnosis must be performed by pediatricians. Therefore, for other departments, the system will block the trigger rule reminders. For example, even if an ophthalmologist issues such a diagnosis, it cannot be reported as an infectious disease diagnosis, and the reporting button will be automatically blocked to reduce the misreporting rate for non-specialists.

Setting the non-monitoring time range for follow-up visits: For patients who have follow-up visits within a certain period, the system can be set to exclude monitoring. For example, for patients with viral hepatitis who need regular medical visits and follow-up visits to obtain medication, the system can be set not to monitor the diagnosis name "viral hepatitis," and the time period can be set as permanent, 3 months after reporting, or 6 months after reporting. The time period can be customized.

Monthly rule maintenance and upgrades: For diseases with extremely high trigger frequency but relatively low actual diagnosis rate, the system will disable the rule. For example, the suspected rule for bacterial dysentery is that the chief complaint includes diarrhea. However, there are many diseases that present with diarrhea, which can interfere with the diagnosis of bacterial dysentery, so the suspected rule will be disabled. The upgraded rule will require clinical diagnosis and confirmation through fecal smear examination, which shows a large number



(a)



(b)

Figure 7. Integration of the infectious disease surveillance and decision support system with the doctor workstation provides real-time reminders during medical record writing (a). Public health personnel can review each alert record and modify the content of the report through the system management interface (b).

of red blood cells, a small number of white blood cells, Sha-Ko-Liden crystals, or visible amoeba trophozoites and/or cysts in the tissue.

Evaluation of system effects. The system effects will be evaluated from three perspectives. First, the effectiveness of various types of decision support reminders, including the overall number of triggers, effectiveness rate, and analysis of invalid situations need to be validated. Second, the usability of the system will be analyzed, including the distribution of time from patient arrival to diagnosis, the distribution of time for timely reporting after system reminders, and the time taken by doctors to fill out the report card. Third, interrupted time series methods³³ will be used to compare the proactive reporting mode of doctors before the system is online with the decrease in missed reports after the system is online.

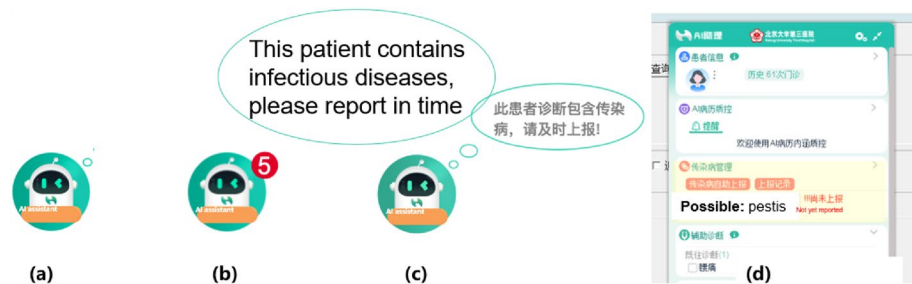


Figure 8. (a) No reminder state is represented by the AI assistant icon. (b) When triggered by a rule, the AI assistant icon displays a numbered reminder in the upper-right corner. (c) The bubble represents a non-interruption-based reminder. (d) The pop-up window represents an interruption-based reminder from the AI assistant, which will be highlighted and flash.

Time period	Total	COVID-19 monitoring positive IGG/IGM/PCR	Suspected	Suspected reported	Clinical diagnosis	Clinical diagnosis reported	Confirmed diagnosis	Confirmed diagnosis reported	Direct diagnosis discovery	Direct diagnosis discovery reported
December 2021	1463	547	1	0	6	0	59	32	850	814
January 2022	2486	160	0	0	2	0	29	18	2295	2261
February 2022	1021	157	0	0	6	0	38	18	820	787
March 2022	1228	353	6	0	17	1	40	17	812	772
April 2022	804	460	14	0	13	0	56	14	261	218
May 2022	613	440	8	0	12	0	26	12	127	102
June 2022	544	378	7	0	11	1	46	27	102	68
July 2022	689	438	16	0	9	0	35	18	191	154
August 2022	1083	553	6	0	7	1	32	22	485	437
September 2022	855	442	9	0	10	0	42	27	352	316
October 2022	645	366	9	0	10	0	33	15	227	201
November 2022	596	307	6	0	6	0	63	35	214	182

Table 1. Analysis of infectious disease alerts by the infectious disease alert system from December 2021 to November 2022.

Disease category	Total alerted cases	Reported cases	Disease category	Total alerted cases	Reported cases
Viral hepatitis	8	5	Brucellosis	1	0
Tuberculosis	51	13	Mumps	3	3
Influenza	2247	2234	SARS	2	0
Hand, foot, and mouth disease	3	3	Gonorrhea	7	5
Syphilis	10	2	Leprosy	1	1
Scarlet fever	1	1	Chickenpox	16	15
Schistosomiasis	2	1	Chlamydia trachomatis infection of reproductive tract	26	21
COVID-19	102	0	condyloma acuminatum	1	0
Genital herpes	2	0	Bacterial dysentery	3	1

Table 2. Alerts and reports by disease category in February 2022.

Ethics approval and consent to participate. The study was approved by the Medical Science Research Ethics Committee of Peking University Third Hospital (serial number: IRB00006761-M2022287). Informed consent from the patients was exempt due to the retrospective nature of the study. The ethics committee (Peking University Third Hospital Medical Science Research Ethics Committee) approved the study abandoned informed consent. We confirm all the experiment protocol for involving humans was in accordance to guidelines of national/international/institutional or Declaration of Helsinki in the manuscript.

Results

Rule trigger and false alarm situation. The system was launched in December 2021 and was fully implemented in a large comprehensive hospital and its three branch campuses in China, monitoring and providing alerts for all outpatient, emergency, and inpatient cases. The system is utilized by 1146 doctors across 30 clinical departments in the hospital. These include common infectious disease departments such as Pediatrics and Respiratory departments as well as screening departments where infectious diseases may be identified. For the retrospective research analysis, we chose a stable period with minimal impact from the COVID-19 pandemic, i.e., December 2021 to November 2022, covering a total of 12 months. December 2022 was excluded from the study because of major changes in China's epidemic prevention policies³⁴, which affected the volume of fever-related visits.

During the retrospective research period, the system analyzed 4,497,091 medical records and triggered 12,027 monitoring and warning messages. Among these warnings, 54.63% (n = 6570) were related to infectious diseases and reported, 43.80% (n = 5269) were considered valid by doctors but not reported, and 1.56% (n = 188) were deemed invalid by doctors and not reported. Within the reported warnings, 87.32% were related to COVID-19 with positive results in IGG, IGM antibodies, and nucleic acid testing. The monthly breakdown values of warning results are listed in Table 1, categorized based on the reminder outcomes.

During the retrospective study, among the 12,027 monitoring alert messages triggered, the gender ratio was 1.108:1, with an almost equal proportion of males and females. The disease category and reporting situation for the month of February 2022, which had the highest monitoring volume, are listed in Table 2. Combining Tables 1 and 2, it was found that the suspected and clinically diagnosed rules had reporting rates of 0% and 2.75%, respectively, resulting in a considerable number of ineffective alerts, and doctors marked them as "not meeting the diagnostic criteria". The suspected rule mainly concerned mumps, with the rule being [{"Symptom Name (Present Illness or Chief Complaint)": "Swollen Parotid Gland"}]. The clinical diagnosis rule was mostly based on medical history or symptom monitoring, and based on the actual alert situation, the diseases for which alerts were raised most frequently were genital herpes, condyloma acuminatum, and hand, foot, and mouth disease. The rule for genital herpes was [{"Symptom Name (Present Illness)": "Skin Blister"} and {"Body Part Name": "Vagina"}], most of which were actually diagnosed as shingles, so it is necessary to optimize the rules jointly with experienced clinicians, further subdivide genital herpes, or choose not to use this rule to trigger alerts.

Of the confirmed diagnoses, 51.10% of the alerts were effectively reported in a timely manner, while the remaining 48.90% were marked as reminders being effective but not requiring reporting in the system, as the patients had already been diagnosed in other hospitals or were just returning for follow-up visits. For the diagnoses directly made by doctors in the medical records as presented in Table 1, the effective reporting rate of doctors was 93.70%, and the remaining unreported cases were mainly related to follow-up prescriptions. Overall, this convincingly demonstrates that the alert system has a high effective rate.

System effectiveness analysis. A total of 2523 confirmed infectious disease cases that were randomly sampled and completed reporting from December 2021 to May 2022 were analyzed with regard to the timeliness of diagnosis, effectiveness of reminders, and efficiency of card filling.

Timely diagnosis of infectious diseases will substantially reduce the risk of nosocomial infections caused by patients waiting for diagnosis. The time at which patients were first seen by doctors at the doctor station was taken as the visit time (T_{visit}), and the time at which the system triggered an alert was taken as the monitoring time ($T_{monitoring}$). Equation (1) was used to calculate T_0 , which represents the interval between the patient's visit and the system's alert for infectious diseases, indicating the timeliness of diagnosis. The overall analysis is presented in Table 3. For reproductive tract chlamydia trachomatis infection, gonorrhoea, and syphilis, the first alert time was more than 24 h, mainly because of the rule that positive indicators in the test results were included, and the test report takes 3–4 days to issue. For some patients with pulmonary tuberculosis, hospitalization is required because of lung shadows and common differential diagnosis of lung cancer, and more tests or pathological results

	Influenza	Hand, foot, and mouth disease	Epidemic parotitis	Genital chlamydia infection	Gonorrhoea	Infectious diarrhea	Tuberculosis	Varicella	Viral hepatitis	Bacterial dysentery	Scarlet fever
Total cases	2375	6	1	41	8	4	11	40	2	5	1
$T_0 \leq 30$ min	1429	4	0	0	0	0	0	24	0	0	0
$30 < T_0 \leq 60$ min	137	0	0	0	0	4	0	3	0	5	0
$60 < T_0 \leq 90$ min	118	0	0	0	0		0	2	0	0	1
$90 < T_0 \leq 120$ min	85	1	1	0	0	0	0	0	0	0	0
$2 < T_0 \leq 24$ h	498	1	0	0	0	0	0	11	2	0	0
$24\text{h} < T_0$	108	0	0	41	8	0	11	0	0	0	0
Average time to diagnosis	74% of influenza cases can be diagnosed within 2 h	57 min	1 h 44 min	4 days 6 h	3 days 23 h	50 min	5 days 4 h	62 min	9 h	45 min	77 min

Table 3. Time interval between patients' visit to the system for monitoring and warning infectious diseases T_0 .

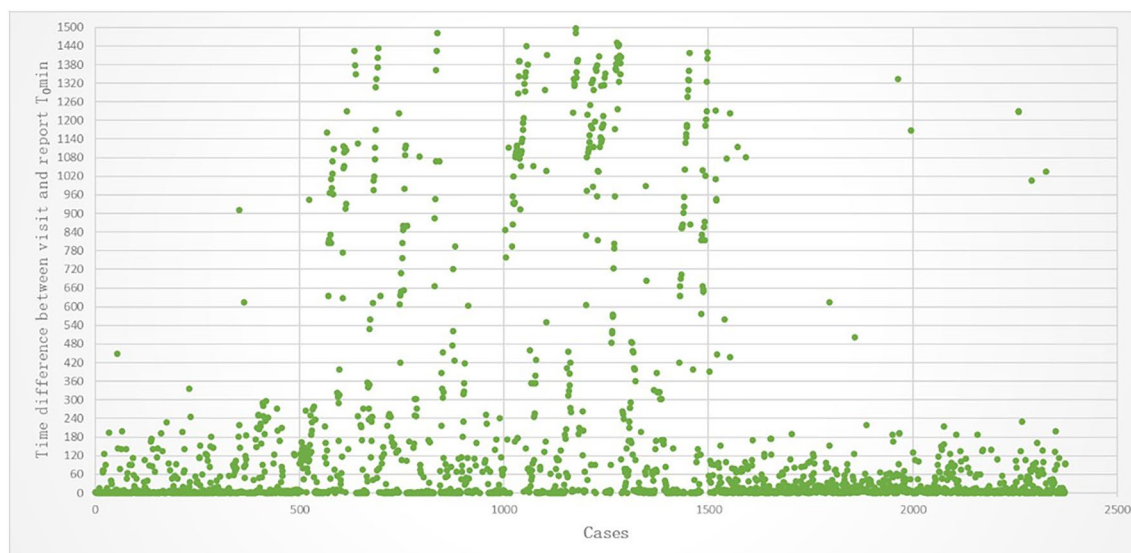


Figure 9. Distribution of time interval between onset and diagnosis for influenza patients.

Reminder and reporting interval T_1	Number of cases	Proportion
$0 < T_1 \leq 1$	180	7.13%
$1 < T_1 \leq 2$	690	27.35%
$2 < T_1 \leq 3$	980	38.84%
$3 < T_1 \leq 4$	304	12.05%
$4 < T_1 \leq 5$	98	3.88%
$5 < T_1 \leq 10$	94	3.73%
$10 < T_1 \leq 20$	39	1.55%
$20 < T_1 \leq 30$	8	0.32%
$30 < T_1 \leq 40$	10	0.40%
$40 < T_1 \leq 50$	7	0.28%
$50 < T_1 \leq 60$	8	0.32%
$60 < T_1 \leq 90$	30	1.19%
$90 < T_1$	75	2.97%

Figure 10. Analysis of time interval between reminder and reporting by doctors.

are needed for confirmation^{35,36}. In addition, pulmonary tuberculosis is sometimes discovered through physical examination, which takes a long time to report. Based on the symptoms of respiratory infectious diseases in the early stages of the SARS epidemic in 2003 and the COVID-19 pandemic in 2019^{37,38}, respiratory infectious diseases are the main cause of nosocomial transmission risk for hospitals. Therefore, the shorter the diagnosis time of respiratory tract diseases, the higher the reduction of nosocomial transmission risk. As shown in Fig. 9, 60% of influenza cases can remind doctors that patients meet the diagnostic criteria for influenza within 30 min. 4.5% of influenza cases have a diagnosis time of more than 24 h, mainly owing to hospitalized patients who were initially diagnosed with a disease other than influenza, then later tested for influenza antigen during their hospital stay, and determined to have influenza. For genital chlamydia infection and gonorrhoea, the average time of T_0 exceeds 24 h, mainly as it requires the support of test results, which take 3 days to obtain. For tuberculosis, 50% of tuberculosis patients are detected through medical examinations, wherein it takes longer to generate reports than regular outpatient and inpatient visits.

$$T_{\text{monitoring}} - T_{\text{visit}} = T_0. \quad (1)$$

Timeliness analysis of reminders is important. The time interval between receiving a reminder and reporting it by doctors is analyzed. The triggering time of the alert in the system is taken as the monitoring time ($T_{\text{monitoring}}$), and the time when the doctor clicks the "report" button is taken as the reporting time (T_{report}). Equation (2) is used to calculate T_1 to obtain the time difference between the reminder and the report, which indicates the timeliness of the reminder. The results are shown in Fig. 10. 89.26% of infectious disease cases can be confirmed

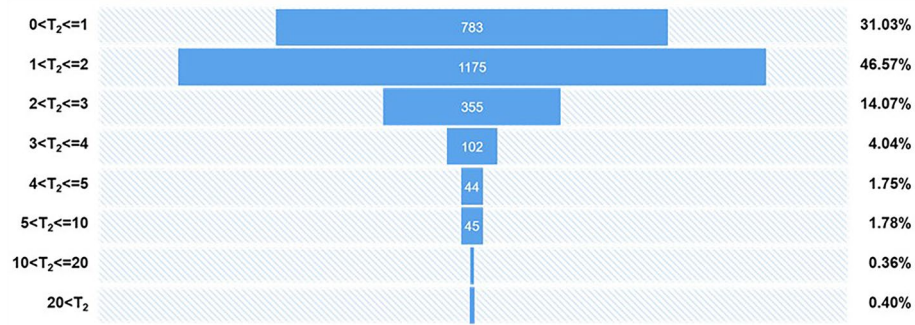


Figure 11. Time for doctors to fill out the report card is mainly concentrated within 2 min (accounting for 77.61%).

and reported within 5 min after the reminder is received. According to the system’s reminder setting, doctors are reminded of positive results of previously treated patients after they complete the current patient’s consultation, so some cases may take more than 5 min to report. For 2.97% of cases that took longer than 90 min, it was mainly due to positive test results being reported during doctor shift changes or holidays, and other doctors had to handle them.

$$T_{report} - T_{monitoring} = T_1 \tag{2}$$

Analysis of efficiency improvement in filling out reporting cards. The Chinese Center for Disease Control and Prevention’s infectious disease reporting cards contains 33 items. Before the system was launched, doctors mainly filled out the form using independent electronic spreadsheets, and patient information and addresses had to be obtained by asking the patient to enter the information. According to Wang et al.³⁹, it took an average of 5 min for a doctor to complete a reporting form. After the system was launched, patient basic information, address, diagnosis, and other information were directly obtained through the system and electronic medical record system, and doctors only needed to verify and check the content related to infectious diseases, such as whether close contacts had the same symptoms after onset, disease classification, and so on. In order to verify the time saved by the system’s automatic docking, the time when the doctor clicks the "report" button is used as the reporting time T_{report} , and the time when the doctor completes filling out the form and clicks the "submit" button is used as $T_{confirm}$. Equation (3) is used to calculate T_2 to obtain the time it takes for doctors to fill out the reporting form. According to Fig. 11, 31.03% of doctors can complete the form in less than 1 min, 46.57% can complete it in 1–2 min, and less than 5% of doctors take more than 5 min, mainly owing to interruptions from seeing patients or answering phone calls while filling out the form, causing the page to stay on the reporting interface, resulting in the reporting time limit getting exceeded. Overall, by automatically filling in information, the average reporting time for each form filled out by a doctor has been reduced from an average of 5 min to less than 2 min, significantly improving the efficiency of doctors’ work.

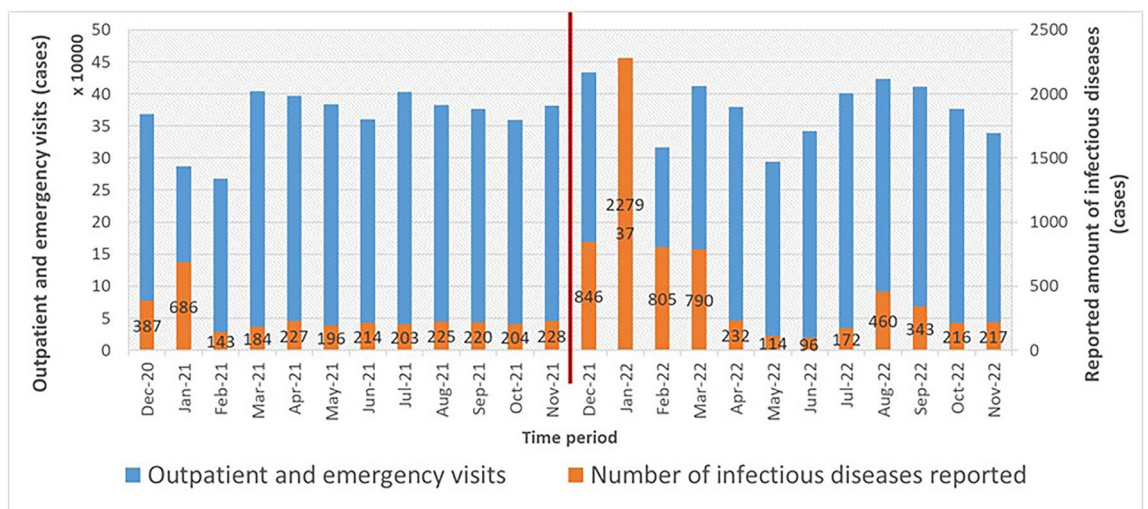


Figure 12. Outpatient and emergency volume and the number of reported infectious diseases before and after the system went online, with a red dividing line indicating the comparison before and after the launch.

$$T_{confirm} - T_{report} = T_2. \quad (3)$$

Using interrupted time series to reduce underreporting rate analysis. In the past, the management of infectious disease reporting mainly relied on manual report verification⁴⁰, and the use of clearly diagnosed report analysis in electronic medical records could remind doctors to report within the prescribed time limit in a timely manner. The main causes of missed reports that occur after the deadline are: (1) delayed laboratory test results; in clinical practice, many test results, such as bacterial culture and syphilis antibody testing, usually take 3–5 days until the results are ready, and in some cases, it may take several weeks for bacterial cultures to yield positive results, which cannot be promptly fed back to doctors; (2) patients transferred from outpatient or emergency departments to hospitals, as doctors usually focus on the primary symptoms of patients for treatment and may neglect to simultaneously check for infectious diseases; (3) infectious diseases such as hepatitis B and tuberculosis can be detected at hospital medical examination centers, but as the time required to issue medical examination reports is longer than that for normal outpatient and emergency departments, there is a risk of missed reports. According to Eq. (4), first given in Chen et al.⁴¹, the underreporting rate $Underreport_0$ is calculated as follows:

$$\frac{\text{Number of under - reported cases}}{\text{Number of infectious disease cases in the same period}} \times 100\% = Underreport_0. \quad (4)$$

Figure 12 shows the trend relationship between the number of hospital outpatient and emergency visits and the number of infectious disease reports from December 2020 to November 2022. Except for a noticeable increase in January 2022, the number of infectious disease reports remained relatively stable in other months. The p-value of the KS test for the underreporting rate data is 0.261 (no significant difference from the normal distribution), indicating that the data follow a normal distribution.

The total underreporting rate in 2021 is $179/3117 = 5.74\%$, and the total underreporting rate in 2022 is $80/6570 = 1.22\%$.

The Interrupted Time Series (ITS) method is used to evaluate the effectiveness of the reduction in underreporting rates after system implementation by analyzing the change in slope and immediate level change at the intervention point. The data and parameters are presented in Table 4. Let Y represent the under-reporting rate and X_1 be the time variable for counting, with $X_1 = 1, 2, 3, \dots, n$. X_2 indicates the intervention, with $X_2 = 0$ before the intervention (December 2020 to November 2021), and $X_2 = 1$ after the intervention (December 2021 to November 2022). $X_3 = 0$ indicates observations before the intervention, and $X_3 = X_1$ indicates observations after the intervention.

Time period	Number of infectious diseases reported	Number of underreported cases	Underreport ₀	X ₁	X ₂	X ₃
December 2020	387	21	5.43	1	0	0
January 2021	686	38	5.54	2	0	0
February 2021	143	8	5.59	3	0	0
March 2021	184	10	5.43	4	0	0
April 2021	227	13	5.73	5	0	0
May 2021	196	12	6.12	6	0	0
June 2021	214	13	6.07	7	0	0
July 2021	203	12	5.91	8	0	0
August 2021	225	13	5.78	9	0	0
September 2021	220	13	5.91	10	0	0
October 2021	204	12	5.88	11	0	0
November 2021	228	14	6.14	12	0	0
December 2021	846	12	2.08	13	1	13
January 2022	2279	24	1.72	14	1	14
February 2022	805	9	1.75	15	1	15
March 2022	790	9	1.05	16	1	16
April 2022	232	4	1.12	17	1	17
May 2022	114	2	1.14	18	1	18
June 2022	96	2	1.42	19	1	19
July 2022	172	2	1.16	20	1	20
August 2022	460	6	1.30	21	1	21
September 2022	343	4	1.17	22	1	22
October 2022	216	3	1.39	23	1	23
November 2022	217	3	1.38	24	1	24

Table 4. Monthly underreporting rate and parameters after and before the system went online.

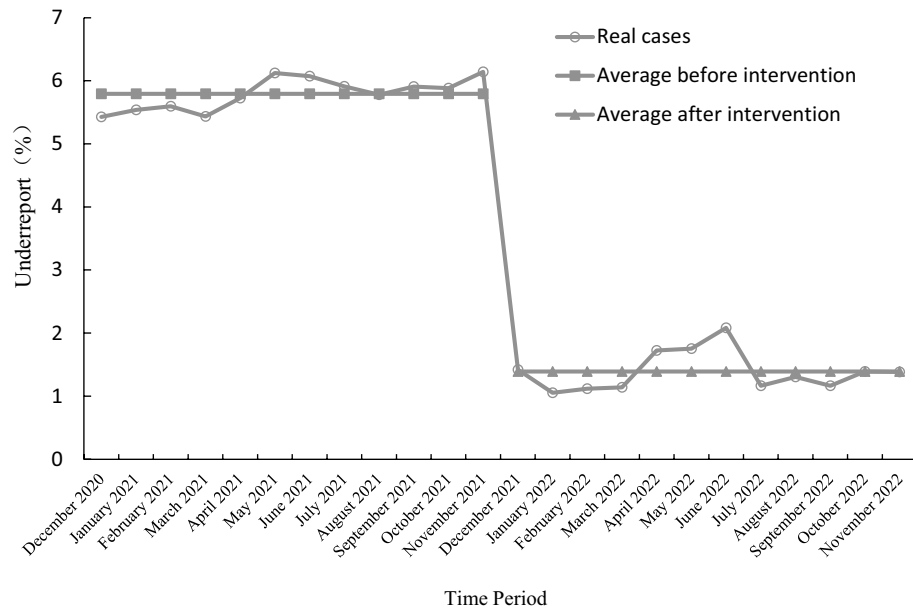


Figure 13. Change model of the level of underreporting rate.

- (1) Underreporting rate before and after the intervention: The average underreporting rate before the intervention was 5.7950%, the average under-reporting rate after the intervention was 1.3913%, and the average under-reporting rate after the intervention was 4.4037% lower than that before the intervention.
- (2) Level change model:

Using underreporting rate as the dependent variable and X_2 as the independent variable, a simple linear regression model is fit:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_2 X_2 = 5.7950 - 4.4037 X_2.$$

When $X_2=0$, the predicted missed report rate before intervention is:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_2 X_2 = 5.7950.$$

When $X_2=1$, the predicted missed report rate after intervention is:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_2 X_2 = 1.3913.$$

The regression coefficient -4.4037 indicates the impact of the intervention on the level of missed report rate, and as the coefficient is negative, the missed report rate after intervention is actually reduced by 4.4037 percentage points compared to before intervention. The p-value of the t-test for the regression coefficient $\hat{\beta}_2$ is 0, <0.01 , indicating that the difference is statistically significant, as shown in Fig. 13.

- (3) Level and slope change model:

The level and slope change model are fit as per Eq. (5):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon. \tag{5}$$

Variables X_1, X_2, X_3 are substituted into Eq. (5). Before the experiment: $X_1=1, 2, \dots, n, X_2 = X_3 = 0$, then the model is:

$$Y = \beta_0 + \beta_1 X_1 + \epsilon$$

After the intervention: $X_2 = 1, X_3 = X_1$, then the model is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon_t = \beta_0 + \beta_1 X_1 + \beta_2 \times 1 + \beta_3 X_1 + \epsilon = \beta_0 + \beta_2 + \beta_1 + \beta_3 X_1 + \epsilon_t = \beta_0^* + \beta_1^* X_1 + \epsilon.$$

β_1 is the slope before intervention, β_2 is the level change, β_3 is the slope change, $\beta_1^* = \beta_1 + \beta_3$ is the slope after intervention, and $\beta_0^* = \beta_0 + \beta_2$ is the intercept after intervention.

Using data from Table 4 and fitting the multiple linear regression, the level and slope change model is obtained as:

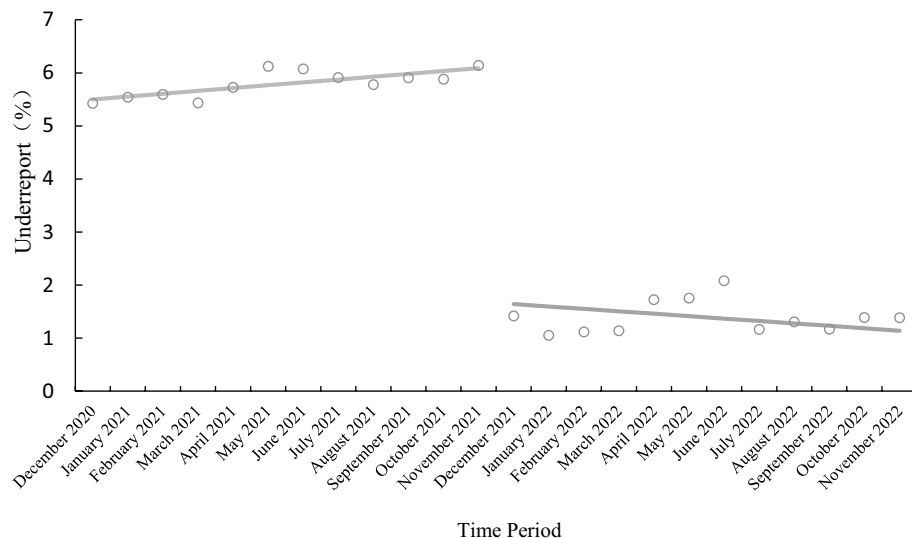


Figure 14. Model of the level and slope change of the missed report rate.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 = 5.4453 + 0.0537X_1 - 3.2092X_2 - 0.0994X_3.$$

The slope before intervention is $\hat{\beta}_1 = 0.0537$, and the slope after intervention is $0.0537 - 0.0994 = -0.0453$. The slope change is $\hat{\beta}_3 = -0.0994$. The p-values of the t-test for $\hat{\beta}_1, \hat{\beta}_2$, and $\hat{\beta}_3$ are 0, 0.0119, and 0.0017, respectively, indicating that there is a significant change in the slope before and after intervention, as shown in Fig. 14.

Discussion

Real case generation association atlas improves the integrity and applicability of rules. Previous studies^{42,43} relied solely on neural network models for early warning prediction and diagnostic recommendations, making it difficult for doctors to rely on AI diagnostic recommendations in real clinical settings. The knowledge- and graph-based rule monitoring proposed in this study provides intuitive warning criteria and evidence-based decision-making support. Knowledge-based decision support does not simply digitize guidelines; rather, it requires retrospective research to continuously optimize the rules. The use of graphs further enhances and visualizes the rules summarized based on clinical guideline knowledge, facilitating rule modification that conforms to doctors' writing habits and extracts symptom keywords. After the system's launch in December 2021, the rules were continuously optimized. For instance, the rules related to influenza involved symptoms such as "fever," "cough," "headache," and "sore throat". However, these symptoms can be expressed differently, and doctors may have various writing habits. Therefore, it was necessary to standardize and unify the synonyms of symptoms. To achieve synonym transformation, the specific method involved converting the words into word vectors, thereby mapping semantically similar words to vectors with close proximity in vector space. This process was achieved by calculating the cosine similarity between the word vector of the target word and all standard word vectors. By finding the standard word vector with the closest distance, the target word was replaced with the corresponding standard word vector, completing the synonym transformation. In addition, in order to protect the privacy of infectious disease patients, doctors usually diagnose "urinary tract infection" rather than "genital chlamydia trachomatis infection," so relevant content is added to the monitoring rules. Based on the infectious disease guidelines and knowledge graph, the rules for assisted warning have authoritative guidance, are supplemented by graph symptoms based on real medical records, meet the hospital's personalized decision support needs, and improve the accuracy of early warning while reducing the possibility of symptom omission.

Necessity of establishing a rule update and maintenance team. The rules in the system are dynamic and require timely updates as clinical guidelines evolve and new treatment options emerge. The rule update and maintenance team includes the following roles:

Medical Experts: Senior clinical experts with advanced titles are responsible for providing updates on the latest clinical guidelines and treatment protocols.

Information Engineers: They assist medical experts in configuring the rules within the system, which primarily involves the configuration of data field names.

Review Experts: They are responsible for reviewing the modified content, working closely with medical experts to address any potential conflicts.

Quality Management Experts: They monitor the updated rules. If a rule is marked as invalid more than 50% of the time within three months, they promptly provide feedback to Medical Experts and Review Experts for rule revision.

The collaboration among these team members ensures the continuous accuracy and relevance of the rules in the system, aligning them with the latest medical knowledge and best practices.

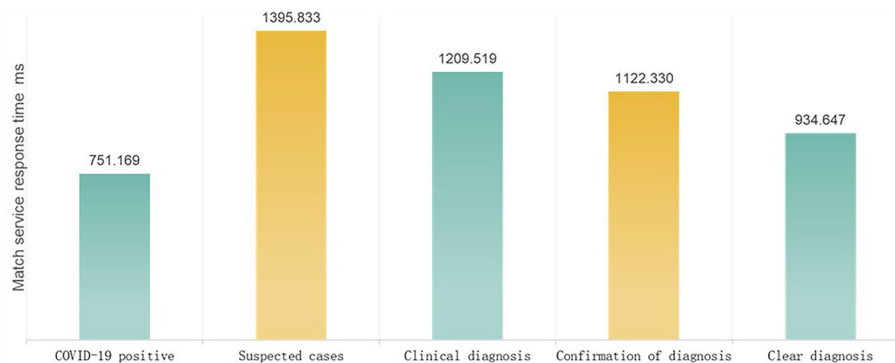


Figure 15. Response efficiency analysis of real-time monitoring service.

The knowledge-based monitoring model proposed in this study can be extended to monitor other diseases.

The comprehensive medical record monitoring method provides effective clinical decision support for non-communicable disease departments, such as the unexplained acute hepatitis in children that occurred in multiple countries in 2022⁴⁴. The system provides pediatricians with early warning of viral hepatitis patients, enabling them to take relevant protective measures in advance. This study analyzed the efficiency of the system's real-time monitoring service response through log monitoring. As shown in Fig. 15, because suspected diagnosis, clinical diagnosis, and confirmed diagnosis all involve word segmentation and matching of medical records, the response time is slightly longer than that of direct diagnosis and positive test calls, but it can still respond within approximately 1 s, meeting the hospital's requirement for real-time disease warning monitoring. In addition, this model can be expanded to monitor other diseases based on rule content, such as foodborne disease monitoring, which will help reduce the incidence of diseases, timely detect outbreaks of microbial foodborne diseases, and quickly respond to control the spread of microbial foodborne diseases^{41,45}. By configuring diagnoses of foodborne diseases such as "pesticide poisoning, pufferfish poisoning, mushroom poisoning, alcohol poisoning, and vibrio parahaemolyticus disease" as monitoring rules, and relevant test rules such as "bacterial culture, [{"test item": "bacterial culture"}] and [{"qualitative test result": "vibrio parahaemolyticus"}]", it can signal the alert as "foodborne monitoring positive."

Limitations. Considering that this study mainly targets only a few specific medical institutions in Beijing, China, where the incidence rates of some infectious diseases are low, the rules for infectious diseases that are affected by environmental exposure, such as echinococcosis, cannot currently be verified for their degree of conformity with actual situations. In addition, as the system adopts the MQ message mechanism⁴⁶ to receive diagnostic data, there may be data transmission failures during peak data periods, which may cause the system to fail to monitor and issue warnings. When applying this system to other electronic health record systems, the differences between various medical institutions must be taken into account. For instance, there may be variations in data formats and programming languages among these systems. Therefore, some modifications are required to ensure seamless integration and compatibility with the diverse systems used in various medical institutions.

Results

This study proposes a knowledge-based infectious disease monitoring and decision support system and validates its effectiveness and reliability. The system can accurately trigger valid alerts and has a low rate of false alarms, improving the infectious disease intelligent decision support for clinical doctors and enhancing the monitoring and management efficiency of disease control departments. In the event of a resurgence of infectious diseases, the rule update team can maintain the rule library based on the latest guidelines. Through the infectious disease monitoring and decision support system, doctors can receive assistance in determining whether a case is infectious, facilitating rapid diagnosis reporting. For hospital management departments, the system's increased reporting capabilities are complemented by a trend prediction model based on our previous research⁴, which enables early warning of outbreaks, thereby enhancing epidemic response capabilities and timeliness.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 10 March 2023; Accepted: 2 August 2023

Published online: 14 August 2023

References

- Jang, H. *et al.* Relationship between economic loss and anxiety during the coronavirus disease 2019 pandemic: Moderating effects of knowledge, gratitude, and perceived stress. *Front. Psychiatry* **13**, 904449 (2022).

2. Chen, X., Gong, W., Wu, X. & Zhao, W. Estimating economic losses caused by COVID-19 under multiple control measure scenarios with a coupled infectious disease-economic model: A case study in Wuhan, China. *Int. J. Environ. Res. Public Health* **18**(22), 11753 (2021).
3. Sakamoto, S., Terashita, D. & Balter, S. Liaison public health nurse project: Innovative public health approach to combat infectious disease in hospitals. *J. Public Health Manag. Pract.* **26**(6), 557–561 (2020).
4. Wang, M., Lee, C., Wang, W., Yang, Y. & Yang, C. Early warning of infectious diseases in hospitals based on multi-self-regression deep neural network. *J. Healthc. Eng.* **2022**, 1–13 (2022).
5. Nestor, C. C., Donnelly, M., Connors, S., Morrison, P. & Boylan, J. Early warning scores and critical care transfer—Patient heterogeneity, low sensitivity, high mortality. *Ir. J. Med. Sci.* **191**(1), 119–126 (2022).
6. Li, W. *et al.* PulseNet China, a model for future laboratory-based bacterial infectious disease surveillance in China. *Front. Med.* **6**(4), 366–375 (2012).
7. Saperstein, Y., Ong, S. Y., Al-Bermani, T., Park, J., Saperstein, Y., Olayinka, J., Jaiman, A., Winer, A., Salifu, M. O., McFarlane, S. I. COVID-19 guidelines changing faster than the virus: Implications of a clinical decision support app. *Int. J. Clin. Res. Trials.* **5**(2) (2020).
8. de Souza, A. A. *et al.* Simple hemogram to support the decision-making of COVID-19 diagnosis using clusters analysis with self-organizing maps neural network. *Soft Comput.* **27**, 3295–3306 (2021).
9. Albin, J. S. *et al.* Development and implementation of a clinical decision support system tool for the evaluation of suspected monkeypox infection. *J. Am. Med. Inform. Assoc.* **29**(12), 2124–2127 (2022).
10. Baumann, L. A., Baker, J. & Elshaug, A. G. The impact of electronic health record systems on clinical documentation times: A systematic review. *Health Policy* **122**(8), 827–836 (2018).
11. dos Santos, H. D. P., Damasio, J. O., Ulbrich, A. H. D. P. S., Vieira, R. Opportunities and challenges in fall risk management using EHRs and artificial intelligence: A systematic review. In *23rd International Conference on Enterprise Information Systems (ICEIS): 2021 Apr 26–28 2021; Electr Network.* 626–633 (2021).
12. Zhang, T. *et al.* Early surveillance and public health emergency responses between novel coronavirus disease 2019 and avian influenza in China: A case-comparison study. *Front. Public Health* **9**, 629295 (2021).
13. Yang, S. *et al.* Epidemiological features of and changes in incidence of infectious diseases in China in the first decade after the SARS outbreak: An observational trend study. *Lancet Infect. Dis.* **17**(7), 716–725 (2017).
14. Walsh, S. *et al.* Decision support systems in oncology. *Jco Clin. Cancer Inform.* **3**, 1–9 (2019).
15. Sutton, R. T. *et al.* An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digit. Med.* **3**(1), 17 (2020).
16. Ronan, C. E., Crable, E. L., Drainoni, M.-L. & Walkey, A. J. The impact of clinical decision support systems on provider behavior in the inpatient setting: A systematic review and meta-analysis. *J. Hosp. Med.* **17**(5), 368–383 (2022).
17. Jing, X. *et al.* Ontologies applied in clinical decision support system rules: Systematic review. *JMIR Med. Inform.* **11**, e43053–e43053 (2023).
18. Dalal, A. K. *et al.* The impact of automated notification on follow-up of actionable tests pending at discharge: A cluster-randomized controlled trial. *J. Gen. Intern. Med.* **33**(7), 1043–1051 (2018).
19. Khan, B. A. *et al.* Clinical decision support system and incidence of delirium in cognitively impaired older adults transferred to intensive care. *Am. J. Crit. Care* **22**(3), 257–262 (2013).
20. Tao, L. *et al.* Accuracy and effects of clinical decision support systems integrated with BMJ best practice-aided diagnosis: Interrupted time series study. *JMIR Med. Inform.* **8**(1), 56–70 (2020).
21. Chien, S. C. *et al.* Alerts in clinical decision support systems (CDSS): A bibliometric review and content analysis. *Healthcare* **10**(4), 601 (2022).
22. Park, H. *et al.* Appropriateness of alerts and physicians' responses with a medication-related clinical decision support system: Retrospective observational study. *JMIR Med. Inform.* **10**(10), e40511 (2022).
23. Binh-Phi, N., Reese, T., Decker, S., Malone, D., Boyce, R. D., Beyan, O. Implementation of clinical decision support services to detect potential drug-drug interaction using clinical quality language. In *17th World Congress of Medical and Health Informatics (MEDINFO): 2019 Aug 25–30 2019; Int Med Informat Assoc, Lyon, FRANCE.* 724–728 (2019).
24. Wang, M., Wei, Z., Jia, M., Chen, L. & Ji, H. Deep learning model for multi-classification of infectious diseases from unstructured electronic medical records. *BMC Med. Inform. Decis. Making* **22**(1), 1–13 (2022).
25. Dissanayake, P. I., Colicchio, T. K. & Cimino, J. J. Using clinical reasoning ontologies to make smarter clinical decision support systems: A systematic review and data synthesis. *J. Am. Med. Inform. Assoc.* **27**(1), 159–174 (2020).
26. Han, C., Li, M., Haihambo, N., Cao, Y. & Zhao, X. Enlightenment on oscillatory properties of 23 class B notifiable infectious diseases in the mainland of China from 2004 to 2020. *PLoS One* **16**(6), e0252803 (2021).
27. Wang, L. *et al.* Health system reform in China 2 Emergence and control of infectious diseases in China. *Lancet* **372**(9649), 1598–1605 (2008).
28. Zhou, J. *et al.* Moving toward a standardized diagnostic statement of pituitary adenoma using an information extraction model: A real-world study based on electronic medical records. *BMC Med. Inform. Decis. Making* **22**(1), 319 (2022).
29. Chen, T., Xu, R., He, Y. & Wang, X. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Syst. Appl.* **72**, 221–230 (2017).
30. Li, L., Hou, L. IEEE: Combined attention mechanism for named entity recognition in chinese electronic medical records. In *7th IEEE International Conference on Healthcare Informatics (ICHI): 2019 Jun 10–13 2019; Xian, PEOPLES R CHINA.* 476–477 (2019).
31. <https://github.com/grapheco/InteractiveGraph/releases/tag/0.3.1>.
32. Chen, B. *et al.* Changes in incidence of notifiable infectious diseases in China under the prevention and control measures of COVID-19. *Front. Public Health* **9**, 728768 (2021).
33. Degli Esposti, M. *et al.* Can synthetic controls improve causal inference in interrupted time series evaluations of public health interventions?. *Int. J. Epidemiol.* **49**(6), 2010–2020 (2020).
34. Wei, L., Liu, N., Chen, J. & Sun, J. Topic evolution of Chinese COVID-19 policies based on co-occurrence clustering network analysis. *Sustainability* **14**(4), 2411 (2022).
35. Chai, M. & Shi, Q. The effect of anti-cancer and anti-tuberculosis treatments in lung cancer patients with active tuberculosis: A retrospective analysis. *BMC Cancer* **20**(1), 1–7 (2020).
36. Lu, Y. *et al.* CT imaging of coexisting pulmonary tuberculosis and lung cancer. *Chin. J. Radiol.* **47**(1), 8–12 (2013).
37. Chen, S. Y. *et al.* Sequential symptomatic analysis in probable severe acute respiratory syndrome cases. *Ann. Emerg. Med.* **43**(1), 27–33 (2004).
38. Bowyer, R. C. E. *et al.* Characterising patterns of COVID-19 and long COVID symptoms: Evidence from nine UK longitudinal studies. *Eur. J. Epidemiol.* **374**, 1648 (2023).
39. Mengying, W. Research of infectious disease early warning system in real scene of hospital-based on big data. *Chin. Hosp. Manag.* **42**(3) (2022).
40. Qiyambo. Underreporting and report quality analysis of statutory infectious diseases in Yunnan Province in 2019. *Mod. Prev. Med.* **48**(8) (2021).
41. Chen, J. *et al.* Practice and consideration of foodborne disease surveillance in Zhejiang Province. *Chin. J. Food Hyg.* **33**(1), 47–52 (2021).

42. Ambika, M., Raghuraman, G. & SaiRamesh, L. Enhanced decision support system to predict and prevent hypertension using computational intelligence techniques. *Soft. Comput.* **24**(17), 13293–13304 (2020).
43. Natalia, K., Mikhail, B., Michael, K., Anna, D., Igor, I. Convolutional neural networks application in cardiovascular decision support systems. In *2nd International Conference on Image and Video Processing, and Artificial Intelligence (IPVAI): 2019 Aug 23–25 2019; Shanghai, Peoples R China.* (2019).
44. Gong, K. *et al.* Acute hepatitis of unknown origin in children: A combination of factors. *Front. Pharmacol.* **13**, 1056385 (2022).
45. Tilashalski, F. P., Sillence, E. M., Newton, A. E. & Biggerstaff, G. K. Enhancing response to foodborne disease outbreaks: Findings of the foodborne diseases centers for outbreak response enhancement (FoodCORE), 2010–2019. *J. Public Health Manag. Pract.* **28**(4), E702–E710 (2022).
46. Deng, Y.-h., Wu, L.-m., Lu, Q.-w. Application and design of data communication mechanism of information systems based on MQ. In *2nd International Conference on Computer and Automation Engineering (ICCAE): 2010 Feb 26–28 2010; Singapore, Singapore.* 451–454 (2010).

Author contributions

All authors have read and approved the manuscript, and each author has participated sufficiently in developing the project and the manuscript. M.Y.W., Z.H.W., and Y.F.S. contributed to the literature review and analysis of the study and drafting the manuscript. M.J. and W.W. contributed to clinical aspects of technical challenges and drafting the manuscript. H.J. oversaw the technical details and analysis of the study.

Funding

This study was supported by the special fund of Beijing Key Clinical Specialty Construction Project; Capital's Funds for Health Improvement and Research, CFH, 2021-1G-4091; and Beijing Natural Science Foundation (No. Z210003).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023