


# Early warning of epidemics: towards a national intelligent syndromic surveillance system (NISS) in China

Peng Jia ,<sup>1,2</sup> Shujuan Yang<sup>2,3</sup>

**To cite:** Jia P, Yang S. Early warning of epidemics: towards a national intelligent syndromic surveillance system (NISS) in China. *BMJ Global Health* 2020;**5**:e002925. doi:10.1136/bmjgh-2020-002925

**Handling editor** Seye Abimbola

PJ and SY contributed equally.

Received 15 May 2020

Revised 27 August 2020

Accepted 29 August 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

<sup>2</sup>International Institute of Spatial Lifecourse Epidemiology (ISLE), Hong Kong, China

<sup>3</sup>West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

## Correspondence to

Dr Shujuan Yang;  
rekyin@126.com and  
Dr Peng Jia;  
jiapengff@hotmail.com

After the SARS pandemic in 2003, an urgent demand for an effective national disease reporting and surveillance system could not be clearer in China. The National Notifiable Disease Reporting System (NNDRS) operated by the Chinese Center for Disease Control and Prevention (CDC), also known as the China Information System for Disease Control and Prevention, was established in 2004 to facilitate the complete and timely reporting of cases during the outbreak. The outbreak of the COVID-19 has further advanced the demand for an intelligent disease reporting system, also known as the national intelligent syndromic surveillance system (NISS),<sup>1</sup> which would be able to analyse these suspected cases on the basis of prior knowledge and real-time information before a disease is confirmed clinically and in the laboratory. By doing so would tackle the epidemic quickly during the outbreak and even forecast the outbreak accurately and robustly at early stages.

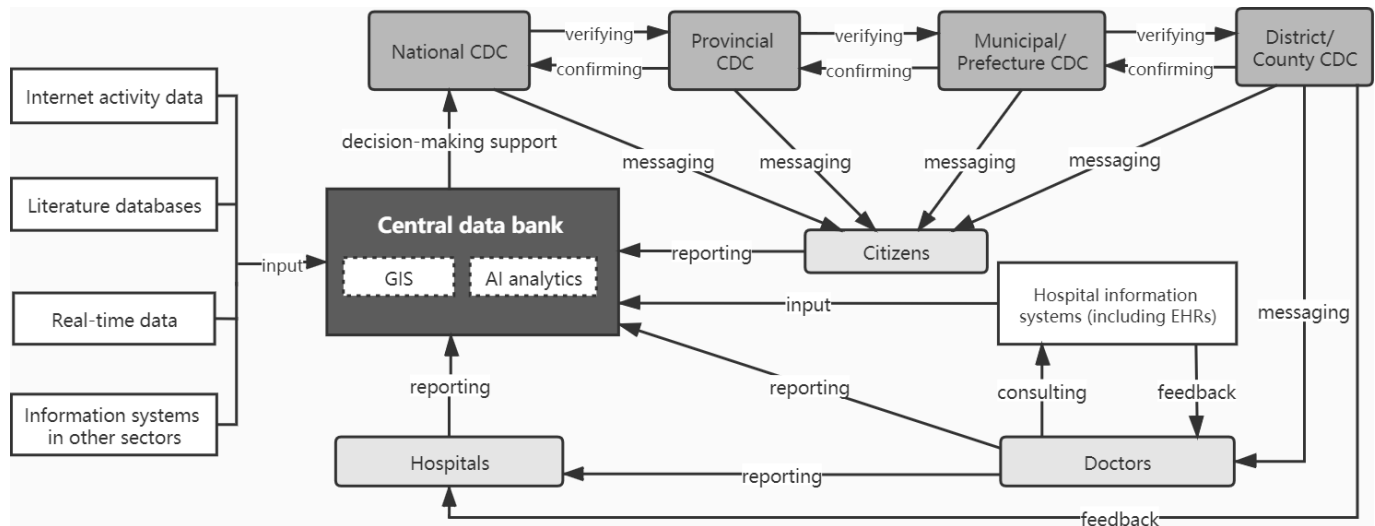
However, it remains difficult to forecast early risks for epidemics in the NISS with only disease cases reported from hospitals. More novel information input from end users and other external sources is required. The current technology enables the end user reporting or input modules in at least seven manners (figure 1). First, the lowest level reporting parties in the current system (ie, hospitals and primary healthcare clinics) should go further down to doctors, who should be able to post and gather the suspected cases they have seen or treated. Doctors see patients directly and are generally more sensitive to suspected cases, which could further reduce the delay under the current NNDRS structure which only allows reporting the potential epidemic from the hospital administration after confirmation by expert panel. This is increasingly important because not only the COVID-19 outbreak added this urgent demand to this system, but the current healthcare reform is shifting more gatekeeper roles down to the

## Summary box

- ▶ A national intelligent syndromic surveillance system (NISS) is necessary in order to tackle the epidemic quickly during the outbreak, and forecast the outbreak accurately and robustly at early stages.
- ▶ Doctors who see patients directly and citizens who are active on the ground are generally more sensitive to and should be enabled to report suspected cases and risk in the NISS.
- ▶ Hospital and other types of information systems (eg, environmental, ecological, agricultural, wildlife and animal) should be tightly linked with the NISS to enable more timely information sharing and make syndromic surveillance.
- ▶ Literature databases containing valuable research findings and knowledge and internet activity data reflecting cyber user awareness should be incorporated into the NISS in a real-time way for warning or fighting the epidemic.
- ▶ Incorporating real-time data into the NISS could greatly facilitate the real-time tracking and consequently guide the epidemic control and prevention work on the ground for curbing the epidemic efficiently.
- ▶ The International Institute of Spatial Lifecourse Epidemiology (ISLE), a global health collaborative research network, has committed to working with multiple stakeholders to codevelop the NISS in China.

primary healthcare clinics. Many western countries are more efficient on this because suspected cases are usually reported first by doctors in private clinics who are the first contact with patients and, under such a structure, are more encouraged to report the suspected cases. With that said, a hierarchical healthcare system is crucial to the successful early detection of infectious diseases. Therefore, efforts on improving the NNDRS should be aligned with the healthcare reform efforts.

Second, citizens should be enabled to report their surrounding risk through this novel crowdsourcing system. Crowdsourcing is a sourcing model in which



**Figure 1** System architecture of a national intelligent syndromic surveillance system in China. AI, artificial intelligence; CDC, Center for Disease Control and Prevention; EHR, electronic health record; GIS, geographic information system. Source: Remade from Nature Medicine.<sup>1</sup>

information can be obtained from a large, relatively open and often rapidly-evolving group of internet users.<sup>2</sup> Such passive surveillance, if well used, could be even more sensitive to the potential risk than reporting by doctors. For example, the crowdsourcing platform *muggenradar* has been used by the Dutch citizens to report the nuisance level of mosquito in order to indicate the potential risk of malaria. Moreover, smartphone-based applications have been popular in China in almost every corner of the daily life, except the disease control and prevention, so crowdsourcing systems should be developed and integrated with the NISSS with high priority. The individuals' communication tools could be even more advanced than the level of information technology infrastructure in some less developed remote areas, where smartphone-based user ends should also be made available for hospitals and doctors to report disease case information, in order to cancel off the low level of information technology infrastructure. Therefore, a crowdsourcing system may work even better in China and other large countries with great variation in economic development. However, citizens should be better educated to be aware of their surrounding risk, which requires incorporating the public health education into the current education system at all stages.

Third, hospital information systems should be tightly linked with the NISSS to enable more timely information sharing and make syndromic surveillance possible.<sup>3</sup> Currently, hospital information systems are not directly linked to the diagnosis-based (or disease-based) NNDRS in China, which could also cause reporting delay and errors (eg, manual typing errors) and should be improved. In the NISSS, information about health events that precede a firm clinical diagnosis should be captured early and rapidly from electronic health records (EHRs), and analysed frequently

to detect signals that might indicate an outbreak requiring investigation. This has been unprecedentedly possible since artificial intelligence (AI) is nowadays used to predict the future disease risk on the basis of EHRs.<sup>4</sup> AI support is also required to link new symptoms with prior knowledge. In addition, information about hospital resources could be incorporated and updated in the NISSS, which will enable the quick reallocation of the limited healthcare resources among hospitals and even among cities and provinces during the outbreak.<sup>5</sup>

Fourth, information systems in other sections should also be linked with the NISSS, so multisource information could be synthesised to maximise clues for the potential risk for epidemics. For example, linking environmental, ecological, agricultural and wildlife and animal information systems that have been continuously monitoring the nature and human-nature interfaces would help to better early detect the appearance of disease cases with a natural origin, such as COVID-19. Such linkage can also alarm other sectors how they could be affected by the epidemic, so they all could make their own strategies to reduce their loss while effectively avoiding the epidemic, and to alleviate other common issues, such as inadequate staffing and funding in those sectors.<sup>6</sup>

Fifth, literature databases, especially in biomedical fields, normally contain valuable research findings and knowledge and should be incorporated into the NISSS in a real-time way for warning or fighting the epidemic. For example, one article published 2 months ahead of the COVID-19 outbreak revealed several first-time detection of parasites in farmed and wild snakes in Wuhan Huanan Seafood Wholesale Market,<sup>7</sup> which could have been an early warning for that region requiring in-depth investigation. Identifying and integrating scientific publications have been

realised on several commercial social networking sites for scientists and researchers to share papers. AI support is also needed to conduct semantic analyses and identifies potential risk from a sea of knowledge, which can further be automatically positioned by spatial technologies and analysed as a whole.

Sixth, data of internet activities can be leveraged as a complementary source to reflect cyber user awareness, understand the epidemiological factors of diseases and imply the disease risk in users' surroundings. For example, internet-based search engine (volume of search keywords) and social media activity data (eg, Twitter messages, WeChat posts) have been associated with the daily numbers of reported human H7N9 cases<sup>8</sup>; Google Flu Trends was a web service that provided estimates of influenza activity by aggregating Google Search queries.<sup>9</sup> Hence, as the internet and social media are increasingly becoming major sources of health information, such internet surveillance is also more important than ever before in public health emergency control and prevention. Such even more passive surveillance should be an additional module in the NISSS for disease surveillance where the frequently searched disease-related keywords may deserve special attention by CDC.

Last but not least, incorporating real-time data into the NISSS, if set up properly ahead of time, could greatly facilitate the real-time tracking and consequently guide the epidemic control and prevention work on the ground for curbing the epidemic efficiently. Such data-sharing mechanisms and infrastructures would also facilitate timely spatial epidemiological research on the basis of individual-level infected cases linked with respective location data from mobile service providers and/or smartphone-based apps without violating confidentiality requirements. As spatial lifecourse epidemiology is capable of capturing the real-time interaction of three dynamic components (hosts, agents and environments),<sup>10</sup> NISSS running on the basis of real-time data would maximise the strengths of spatial lifecourse epidemiology in the real world, enabling it to outpace the epidemics and realise 'precision epidemic prevention and control'.<sup>10 11</sup> In addition, by setting up such infrastructure ahead of time, the safety of individual confidentiality and information exchange will never be compromised in this powerful system.<sup>11 12</sup>

Some practical aspects of implementation include the integration of disparate data sources in the NISSS and the governance and privacy concerns of the NISSS. The integration of data sharing between agencies can be realised by creating and using an application programming interface (API) for each service, which usually defines many items including the types of calls or requests that can be made and how to make them by data users, and the data formats to use and the conventions to follow by data owners. Data users can request raw data to make forecasts on local machines or servers, which will need data-masking methods (eg, K-Anonymity, L-Diversity, T-Closeness) for better privacy protection. Some non-technical factors may hinder the realisation of raw data sharing, which may require the identification of

a third-party governmental agency and legislation to facilitate data-sharing among sectors; for example, the ongoing Beijing Big Data Action Plan, by Beijing Municipal Bureau of Economy and Information Technology and Beijing Municipal Bureau of Big Data Management, has linked data from government information systems in more than 40 municipal departments in Beijing. With sufficient prior knowledge (eg, knowing which variables are necessary to be used), data users can also request a subset of raw data or processed data to make forecasts, which would decrease the demand for local machine or server configuration and help overcome those non-technical barriers to some extent. However, extra costs will be incurred by creating APIs for (multiple) services, which may require higher-level coordination for cost sharing and/or a data-sharing subsidy among sectors. Lessons could be drawn from some examples of data sharing in other areas but adopting similar approaches. For example, medical big data sharing is gradually being allowed in South Korea; the four main health maintenance organisations in Israel and their affiliated hospitals have used the same electronic medical record (EMR) platform for the past two decades, with access to patient records available to each point of care as needed, and 98% of the population has been using the same linked EMR system for decades.

In addition, multiple stakeholders at different levels of context should sit together, adopting participatory survey and discussion methods to identify more context-specific difficulties and solutions related to the practicality of the NISSS, such as (1) a hierarchy of data sources with the levels of confidentiality and necessity evaluated for each source, (2) more approaches of organising and integrating disparate data sources and conducting analyses, (3) more conceptual, architectural and analytical challenges that would arise and (4) the corresponding solutions that would function best among multiple stakeholders nationally and internationally, including applicability and adaptability of the solutions to similar problems in other countries.

The International Institute of Spatial Lifecourse Epidemiology (ISLE), established as a global health collaborative research network, has committed to identifying the key research issues and priorities for spatial lifecourse epidemiology,<sup>13</sup> advancing the use of state-of-the-art technologies in lifecourse epidemiological research and emerging infectious disease research,<sup>14</sup> and facilitating the quality of reporting of transdisciplinary health research.<sup>15</sup> Establishing the NISSS will be one of the top 10 public health priorities in the next decade, and also on top of ISLE's agenda. Owing to a diverse variety of scholars' backgrounds, ISLE communicates cross-interdisciplinary knowledge and research findings in a plain language with scholars from various disciplines and multiple stakeholders including policy-makers. ISLE has committed to working with multiple stakeholders, from different levels of CDCs to industrial partners, doctors and citizens, to codevelop the NISSS in China. This effort will exemplify the next-generation infectious disease reporting and surveillance system in the 21st century and serve as a model for many other countries in the world.

**Contributors** Both authors have equally contributed to the planning, conduct and reporting of the work described in the article.

**Funding** We thank the National Natural Science Foundation of China (81703279), the Special Funds for Prevention and Control of COVID-19 of Sichuan University (2020scunCoV10010), and the International Institute of Spatial Lifecourse Epidemiology (ISLE) for the research support.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Peng Jia <http://orcid.org/0000-0003-0110-3637>

## REFERENCES

- Jia P, Yang S. China needs a national intelligent syndromic surveillance system. *Nat Med* 2020;26:990.
- Goodchild MF. Citizens as sensors: the world of volunteered geography. *GeoJournal* 2007;69:211–21.
- Henning KJ. What is syndromic surveillance? *MMWR Suppl* 2004;53:5–11.
- Miotto R, Li L, Kidd BA, et al. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep* 2016;6:26094.
- Zarocostas J. What next for the coronavirus response? *Lancet* 2020;395:401.
- Yang J, Siri JG, Remais JV, et al. The tsinghua-lancet Commission on healthy cities in China: unlocking the power of cities for a healthy China. *Lancet* 2018;391:2140–84.
- Xiao X, Qi R, Han H-J, et al. Molecular identification and phylogenetic analysis of *Cryptosporidium*, *Hepatozoon* and *Spirometra* in snakes from central China. *Int J Parasitol Parasites Wildl* 2019;10:274–80.
- Gu H, Chen B, Zhu H, et al. Importance of Internet surveillance in public health emergency control and prevention: evidence from a digital epidemiologic study during avian influenza A H7N9 outbreaks. *J Med Internet Res* 2014;16:e20.
- Dugas AF, Jalalpour M, Gel Y, et al. Influenza forecasting with Google flu trends. *PLoS One* 2013;8:e56176.
- Jia P, Dong W, Yang S, et al. Spatial lifecourse epidemiology and infectious disease research. *Trends Parasitol* 2020;36:235–8.
- Jia P, Yang S. Are we ready for a new era of high-impact and high-frequency epidemics? *Nature* 2020;580:321.
- Jia P. Integrating kindergartener-specific questionnaires with citizen science to improve child health. *Front Public Health* 2018;6:236.
- Jia P, Lakerveld J, Wu J, et al. Top 10 research priorities in spatial lifecourse epidemiology. *Environ Health Perspect* 2019;127:74501.
- Jia P. Spatial lifecourse epidemiology. *Lancet Planet Health* 2019;3:e57–9.
- Jia P, Yu C, Remais JV, et al. Spatial lifecourse epidemiology reporting standards (ISLE-ReSt) statement. *Health Place* 2020;61:102243.