ISSN 2096-7071 (Print)
ISSN 2096-3101 (Online)
CN 10-1629/R1

CHINA CDC WEEKLY

Vol. 7 No. 41 Oct. 10, 2025

CCDC weekly

中国疾病预防控制中心周报

ISSN 2096-7071
9 772096 707255 41

China CDC Weekly

Cover Image: credited from the designer Jiaying Wang.

Perspectives

Artificial Intelligence in Health and Medicine: Progress, Challenges, and Recommendations

Xi Li[1]; Jue Liu[1,#]

ABSTRACT

Artificial intelligence (AI) has broadly reshaped health and medicine, benefiting clinicians, patients, and health systems. However, technical, regulatory, and ethical challenges exist in the application of medical AI, ranging from data scarcity to fairness. We provide our perspective on how to address the major challenges facing widespread clinical adoption from both technical (e.g., building high-quality datasets, using larger and more diverse datasets for training, creating problem formulations that go beyond supervised learning, and combining human skills with AI tools) and ethical (e.g., using highly secure data platforms and strengthening governmental legislation) perspectives.

Artificial intelligence (AI) refers to "machines that mimic cognitive functions similar to the human mind such as learning and problem solving" (*1*). The promise of AI in health and medicine dates back to the 1960s, when ELIZA, one of the first chatbots, was developed to engage in simple dialogue with users (*2*). Owing to rapid advances in AI, an increasing interest in applying AI to medicine has emerged in the last decade (*3*–*4*). AI technology has achieved remarkable success in helping clinicians diagnose and treat diseases, enabling patients to process their own data to improve health, and optimizing health service delivery (*3*). Despite advances, healthcare is below average in the adoption of AI, compared with other industries (*5*), potentially owing to challenges such as bias, privacy and security, and lack of transparency (*6*–*8*). Recommendations have been proposed in the end for the use of AI in health and medicine to overcome these challenges.

PROGRESS IN MEDICAL AI

Only a few years after the first landmark demonstrations of AI algorithms that could detect diseases from medical images, the landscape of medical AI has matured considerably (*4*). AI-enabled medical practice transforms clinical care in terms of medical knowledge, patient journey, and clinical practice (Figure 1). AI systems have been deployed in a wide range of clinical settings that benefit clinicians, patients, and health systems.

AI FOR CLINICIANS

AI has been widely used among clinicians, ranging from specialty doctors to paramedics (*3*,*8*–*10*). This involves pattern recognition using deep neural networks (DNNs) that can help interpret, for instance, vital signs, electrocardiograms, medical scans, endoscopy, pathology slides, skin lesions, and retinal images (*3*). AI systems can serve as medical knowledge resources for clinicians (*9*), simplify clinical text summarization (*10*), and assist in decision-making and guideline adherence (*9*).

Radiology has garnered significant attention in the field of medical AI (*3*). AI systems have been widely used in various medical scans (*3*), including computed tomography scans for lung nodules, pancreatic cancer, and liver masses; echocardiograms; brain scans for evidence of hemorrhage, head trauma, and acute referrals; magnetic resonance imaging; mammography; and bone films for fractures and estimation of aging. AI has repeatedly shown to not only improve the accuracy and efficiency of disease diagnosis but also make more accurate risk predictions and guide treatment (*11*–*13*). For instance, Wang et al. proposed a triage-driven Chinese Lung Nodules Reporting and Data System and found that it achieved a state-of-the-art performance with an area under the curve (AUC) of 0.918 on an internal testing dataset, outperforming the single-dimensional approach (AUC=0.881) (*14*).

Owing to remarkable success in image classification, AI algorithms have then been applied in specialties that heavily rely on the interpretation of medical images

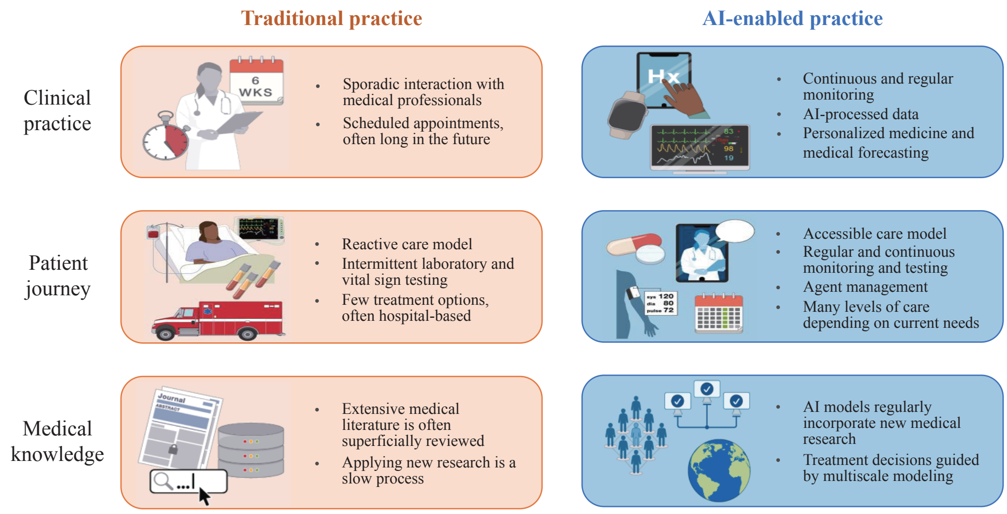


FIGURE 1. Comparison between traditional practice and AI-enabled practice in clinical care.
Abbreviation: AI=artificial intelligence.

such as pathology, ophthalmology, dermatology, cardiology, and gastroenterology (*3–4*). In the field of pathology, using whole-slide imaging, deep learning has made major strides in accelerating workflows for disease diagnosis (*15*) and providing new disease insights (*4*). DNNs have also made progress in gastroenterology, particularly in improving the sensitivity of colonoscopy (*16*).

AI FOR PATIENTS

AI-driven tools can provide patients with accessible conversational interfaces to interact with general medical information and their individual health data in electronic health records (*17*), thereby offering more accessible, proactive, and personalized healthcare (*9*). Medical AI has been employed throughout the human lifespan (Figure 2).

AI algorithms are poised to provide timely and personalized feedback through continuous monitoring (*9*). AI-enabled wearable sensors have shown the ability to identify high-risk individuals (*18*) with high diagnostic accuracy (*19*). Hence, patients are empowered to manage their health and wellbeing in a timely manner (*9*). Furthermore, AI systems promise to revolutionize medical forecasting with multiscale capacity (*9*). Due to the pattern-matching strengths of computers and large data input, forecasting can be applied to the entire context of health, from the molecular level to the cellular, cell/tissue, individual, population, and global levels (*20*). For instance, Li et al. proposed a framework that leverages large language models (LLMs) to estimate individual aging using only health examination reports and found that LLM-predicted ages had superior performance in predicting aging-related diseases than other aging proxies, such as telomere length and frailty index (*21*). The use of AI has enabled more accessible, proactive, and individualized healthcare.

In addition to continuous monitoring, advanced medical screening, and multiscale medical forecasting, the use of AI algorithms has been described in mental health settings. Machine learning can help characterize depression, predict suicide, predict bouts of psychosis in patients with schizophrenia, and predict successful antidepressant medication use through the digital tracking of depression and mood via keyboard interaction, speech, voice, facial recognition, sensors, and interactive chatbots (*3*). In particular, mental health chatbots have demonstrated great potential for reducing the stigma about mental health care and increasing referral rates (*22*).

AI FOR HEALTH SYSTEMS

In addition to these applications, AI algorithms have been integrated into healthcare delivery, making it more precise and efficient (*3*,*9*).

AI-powered scribes have been reported to record patient histories, generate medical notes, handle pre-authorization requests for medications or tests, schedule follow-up appointments, and manage laboratory test results and scans (*9–10*), thereby lowering the potential for medical errors. Moreover, powerful AI models enable earlier diagnoses and interventions, which can improve health outcomes and

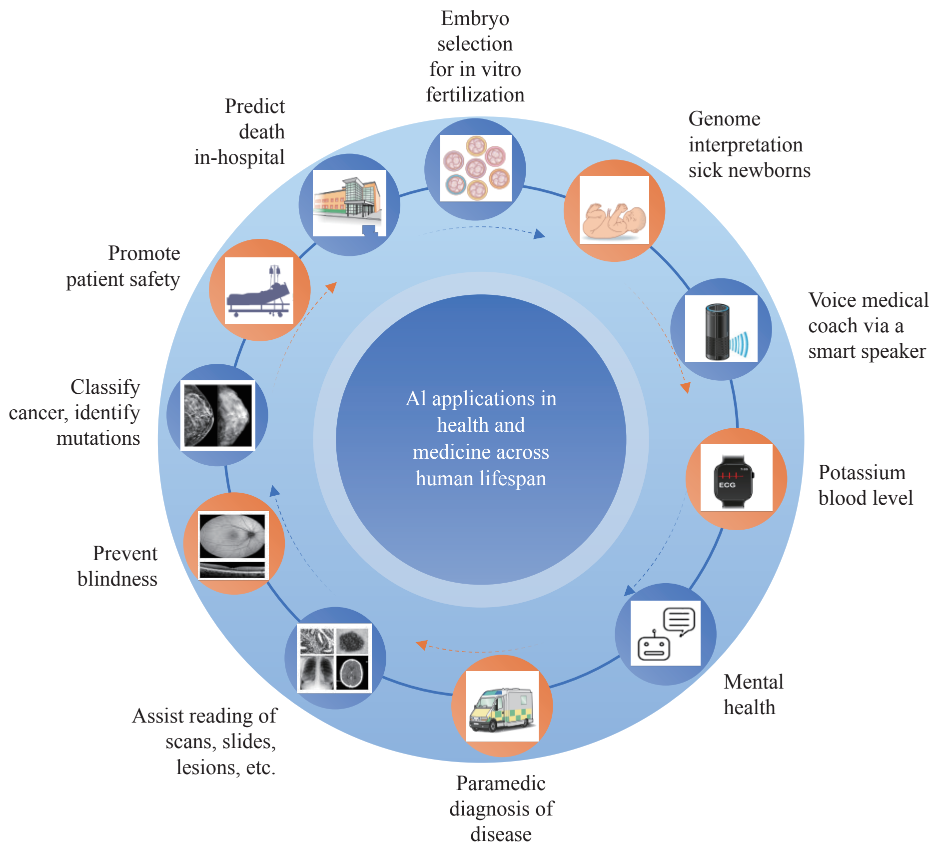


FIGURE 2. Examples of the breadth of AI applications in health and medicine across human lifespan.
Abbreviation: AI=artificial intelligence.

reduce the burden on health systems (*9*). For instance, Du et al. introduced and applied PandemicLLM, a framework with multimodal LLMs that reformulates real-time forecasting of disease spread as a text-reasoning problem to the coronavirus disease 2019 (COVID-19) pandemic, and observed that PandemicLLM performed better in forecasting one-week and three-week hospitalization trends compared with the CDC ensemble model (*23*). In addition to infectious disease prevention and control, AI has been used for immunization. For instance, Hou et al. assessed the effectiveness of a vaccine chatbot in improving human papillomavirus (HPV) vaccination among female middle school students aged 12–15 years and observed that 7.1% of the intervention group had received or scheduled appointments for the HPV vaccine for their daughters versus 1.8% of the control group over a two-week intervention period (*24*).

CHALLENGES IN MEDICAL AI

Notwithstanding striking advances, medical AI faces major technical challenges, particularly in composing training datasets and building model trust (*3–4*). Questions also remain on the regulation of AI in health and medicine and shifts in responsibility throughout the health system (*3–4*). Ethical concerns regarding data privacy, security, and fairness also exist (*3–4*).

Various qualities are desired in AI systems, including reliability, convenience, and ease of integration into clinical workflows (*25*). However, the usefulness of AI relies heavily on how humans provide inputs and interpret the outputs (*4*). Medical data used to train AI systems are limited and possibly erroneous, and the devices required to obtain the input for AI are costly (*4*), impeding both the data collection and deployment of AI systems. Bias in the training data is likely to limit output accuracy and cause major harm to patients (*26*). Numerous AI models now function as uninterpretable black boxes (*3–4*). The data used for training, model performance on any task or population, and variation of the output according to the nondeterministic nature of AI remain unclear (*9*). Accordingly, explaining the predictions of AI models poses a serious technical challenge and their reliability in novel clinical settings remains unclear.

Medical AI poses additional challenges from a regulatory perspective. Regulatory challenges may arise from continual learning (AI models learn from new data over time and adjust to shifts in populations) (*27*), which may contribute to overwriting previously learned patterns or generating new errors (*4*). Moreover, as medical AI takes on more responsibilities, clinicians may become overly reliant on AI systems, leading to a gradual decline in their skills or personal connections with patients (*4*). Hence, a serious concern is that if AI makes mistakes, it remains unclear whether developers, regulators, sellers, or health providers should be held accountable (*4*).

Even further, ethical concerns also exist regarding the application of AI in health and medicine. An overriding issue in medical AI is how well data privacy and security can be assured (*3*,*9*). Bad actors interested in identity theft and other misconduct may take advantage of medical datasets (*4*) containing sensitive information about real patients (*3*). An individual's identity is likely determined by facial recognition or genomic sequences from large datasets (*3*). A risk of deliberate hacking of AI algorithms to harm people on a large scale also exists (*3*). Furthermore, although AI can make healthcare more accessible to underrepresented groups, the disproportionate use of AI systems may reinforce existing inequities in health (*3*–*4*). Intertwined with the concern of widening the present gap in health outcomes is the embedded bias in numerous AI algorithms owing to the lack of inclusion of minorities in datasets (*3*).

RECOMMENDATIONS FOR MEDICAL AI

To build trustworthy AI tools for health and medicine, their design, development, evaluation, and deployment should follow FUTURE-AI, an international consensus guideline based on six guiding principles: fairness, universality, traceability, usability, robustness, and explainability (*28*). Moreover, the applications and optimizations of AI in health and medicine should be based on the complex adaptive systems (CAS) theoretical framework (*29*); that is, AI should adjust its behavior through multi-agent interactions between AI, patients, clinicians, public health practitioners, the public, and decision-makers to achieve continuous learning and adaptive optimization and meet the complex health needs of the population. In practice, in terms of technical challenges, it is necessary to advance AI algorithms by building high-quality datasets through implementing data standardization protocols, multi-institutional collaboration frameworks, and annotation quality control measures; using larger and more diverse datasets for training, such as text, chemical, and genomic sequences (*4*,*9*); creating problem formulations that go beyond supervised learning (*4*); and combining human skills with AI tools (*4*,*30*). Furthermore, new models of health data ownership with rights to individuals, use of highly secure data platforms, and strengthening of governmental legislation are warranted to counter regulatory and ethical issues in medical AI (*3*).

CONCLUSION

The use of AI in health and medicine has made substantial progress: for clinicians, predominantly via rapid and accurate image interpretation and decision-making; for patients, by enabling them to access their own data to promote health; and for health systems, mainly by optimizing healthcare delivery. Despite the major advances in medical AI, challenges related to accuracy, bias, privacy, and ethics persist. These challenges must be addressed before medical AI can be realized to make health systems more accurate, efficient, and accessible to the public worldwide.

Conflicts of interest: No conflicts of interest.

Funding: Supported by National Natural Science Foundation of China (72474005, 72334004); Youth Beijing Scholars Program (087); Science and Technology Innovation 2030—Major Project on "New Generation Artificial Intelligence" (2021ZD0114101).

doi: 10.46234/ccdcw2025.218

Corresponding author: Jue Liu, jueliu@bjmu.edu.cn.

[1] Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China.

Submitted: August 16, 2025
Accepted: September 12, 2025
Issued: September 19, 2025

REFERENCES

1. Sandeep Ganesh G, Kolusu AS, Prasad K, Samudrala PK, Nemmani KVS. Advancing health care via artificial intelligence: from concept to

clinic. Eur J Pharmacol 2022;934:175320. https://doi.org/10.1016/j.ejphar.2022.175320.
2. Weizenbaum J. ELIZA–a computer program for the study of natural language communication between man and machine. Commun ACM 1966;9(1):36 – 45. https://doi.org/10.1145/365153.365168.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44 – 56. https://doi.org/10.1038/s41591-018-0300-7.
4. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022;28(1):31 – 8. https://doi.org/10.1038/s41591-021-01614-0.
5. World Economic Forum. The future of AI-enabled health: leading the way report. Geneva: World Economic Forum; 2025. https://www.weforum.org/publications/industries-in-the-intelligent-age-white-paper-series/healthcare/.
6. Celi LA, Cellini J, Charpignon ML, Dee EC, Dernoncourt F, Eber R, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities–a global review. PLoS Digit Health 2022;1(3):e0000022. https://doi.org/10.1371/journal.pdig.0000022.
7. He JX, Baxter SL, Xu J, Xu JM, Zhou XT, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019;25(1):30 – 6. https://doi.org/10.1038/s41591-018-0307-0.
8. Haibe-Kains B, Adam GA, Hosny A, Khodakarami F, Massive Analysis Quality Control (MAQC) Society Board of Directors, Waldron L, et al. Transparency and reproducibility in artificial intelligence. Nature 2020;586(7829):E14 – 6. https://doi.org/10.1038/s41586-020-2766-y.
9. Fahrner LJ, Chen E, Topol E, Rajpurkar P. The generative era of medical AI. Cell 2025;188(14):3648 – 60. https://doi.org/10.1016/j.cell.2025.05.018.
10. Van Veen D, Van Uden C, Blankemeier L, Delbrouck JB, Aali A, Bluethgen C, et al. Adapted large language models can outperform medical experts in clinical text summarization. Nat Med 2024;30(4):1134 – 42. https://doi.org/10.1038/s41591-024-02855-5.
11. Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. Cancer Cell 2021;39(7):916 – 27. https://doi.org/10.1016/j.ccell.2021.04.002.
12. Gaube S, Suresh H, Raue M, Merritt A, Berkowitz SJ, Lermer E, et al. Do as AI say: susceptibility in deployment of clinical decision-aids. npj Digit Med 2021;4(1):31. https://doi.org/10.1038/s41746-021-00385-9.
13. Huynh E, Hosny A, Guthier C, Bitterman DS, Petit SF, Haas-Kogan DA, et al. Artificial intelligence in radiation oncology. Nat Rev Clin Oncol 2020;17(12):771 – 81. https://doi.org/10.1038/s41571-020-0417-8.
14. Wang CD, Shao J, He YC, Wu JJ, Liu XT, Yang LQ, et al. Data-driven risk stratification and precision management of pulmonary nodules detected on chest computed tomography. Nat Med 2024;30(11):3184 – 95. https://doi.org/10.1038/s41591-024-03211-3.
15. Jackson HW, Fischer JR, Zanotelli VRT, Ali HR, Mechera R, Soysal SD, et al. The single-cell pathology landscape of breast cancer. Nature 2020;578(7796):615 – 20. https://doi.org/10.1038/s41586-019-1876-x.
16. Gong DX, Wu LL, Zhang J, Mu GG, Shen L, Liu J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol 2020;5(4):352 – 61. https://doi.org/10.1016/S2468-1253(19)30413-3.
17. Bernstein IA, Zhang YV, Govil D, Majid I, Chang RT, Sun Y, et al. Comparison of ophthalmologist and large language model chatbot responses to online patient eye care questions. JAMA Netw Open 2023;6(8):e2330320. https://doi.org/10.1001/jamanetworkopen.2023.30320.
18. Attia ZI, Harmon DM, Dugan J, Manka L, Lopez-Jimenez F, Lerman A, et al. Prospective evaluation of smartwatch-enabled detection of left ventricular dysfunction. Nat Med 2022;28(12):2497 – 503. https://doi.org/10.1038/s41591-022-02053-1.
19. Menzies SW, Sinz C, Menzies M, Lo SN, Yolland W, Lingohr J, et al. Comparison of humans versus mobile phone-powered artificial intelligence for the diagnosis and management of pigmented skin cancer in secondary care: a multicentre, prospective, diagnostic, clinical trial. Lancet Digit Health 2023;5(10):e679 – 91. https://doi.org/10.1016/S2589-7500(23)00130-9.
20. Topol EJ. Medical forecasting. Science 2024;384(6698):eadp7977. https://doi.org/10.1126/science.adp7977.
21. Li YJ, Huang Q, Jiang J, Du XS, Xiang WX, Zhang SQ, et al. Large language model-based biological age prediction in large-scale populations. Nat Med 2025. http://dx.doi.org/10.1038/s41591-025-03856-8.
22. Habicht J, Viswanathan S, Carrington B, Hauser TU, Harper R, Rollwage M. Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. Nat Med 2024;30(2):595 – 602. https://doi.org/10.1038/s41591-023-02766-x.
23. Du HR, Zhao Y, Zhao JN, Xu SC, Lin XH, Chen YR, et al. Advancing real-time infectious disease forecasting using large language models. Nat Comput Sci 2025;5(6):467 – 80. https://doi.org/10.1038/s43588-025-00798-6.
24. Hou ZY, Wu ZD, Qu ZQ, Gong LB, Peng H, Jit M, et al. A vaccine chatbot intervention for parents to improve HPV vaccination uptake among middle school girls: a cluster randomized trial. Nat Med 2025;31(6):1855 – 62. https://doi.org/10.1038/s41591-025-03618-6.
25. Cutillo CM, Sharma KR, Foschini L, Kundu S, Mackintosh M, Mandl KD, et al. Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency. npj Digit Med 2020;3:47. https://doi.org/10.1038/s41746-020-0254-2.
26. How to support the transition to AI-powered healthcare. Nat Med 2024;30(3):609-10. http://dx.doi.org/10.1038/s41591-024-02897-9.
27. Lee CS, Lee AY. Clinical applications of continual learning machine learning. Lancet Digit Health 2020;2(6):e279 – 81. https://doi.org/10.1016/S2589-7500(20)30102-3.
28. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ 2025;388:e081554. https://doi.org/10.1136/bmj-2024-081554.
29. Holland JH. Studying complex adaptive systems. J Syst Sci Complex 2006;19(1):1 – 8. https://doi.org/10.1007/s11424-006-0001-z.
30. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. JAMA 2023;330(9):866 – 9. https://doi.org/10.1001/jama.2023.14217.

Jue Liu
School of Public Health, Peking University, Beijing, China

Preplanned Studies

Evaluating Large Language Models' Potential in Field Epidemiology Investigation Based on Chinese Context — Zhejiang Province, China, 2025

Tao Zhang[1,&]; Qifeng Zhao[2,&]; Yaxin Dai[3]; Mengna Wu[1]; Yujia Zhai[1]; Le Xu[1]; Xue Gu[1]; Junfen Lin[1]; Chen Wu[1,#]

Summary

What is already known about this topic?

Large language models (LLMs) have demonstrated considerable potential in clinical applications. However, their performance in field epidemiology, particularly within Chinese-language contexts, remains largely unexplored.

What is added by this report?

This study evaluates six leading LLMs (ChatGPT-o4-mini-high, ChatGPT-4o, DeepSeek-R1, DeepSeek-V3, Qwen3-235B-A22B, and Qwen2.5-Max) using examination questions from the Zhejiang Field Epidemiology Training Program. For multiple-choice questions, all models except DeepSeek-V3 scored below the 75th percentile of junior field epidemiologists, while for case-based questions, LLMs generally outperformed that percentile. However, LLMs demonstrated significant limitations when addressing questions requiring specialized knowledge. Notably, LLMs may generate inaccurate or fabricated references, presenting substantial risks for inexperienced practitioners.

What are the implications for public health practice?

LLMs demonstrate promising potential for supporting epidemiological investigations. Nevertheless, current LLMs cannot replace human expertise in field epidemiology. Their practical implementation faces considerable challenges, including ensuring output accuracy and reliability. Future efforts should prioritize optimizing performance through verified knowledge databases and establishing robust regulatory frameworks to enhance their effectiveness in public health applications.

ABSTRACT

Introduction Large language models (LLMs) have demonstrated potential applications across diverse fields, yet their effectiveness in supporting field epidemiology investigations remains uncertain.

Methods We assessed six prominent LLMs (ChatGPT-o4-mini-high, ChatGPT-4o, DeepSeek-R1, DeepSeek-V3, Qwen3-235B-A22B, and Qwen2.5-max) using multiple-choice and case-based questions from the 2025 Zhejiang Field Epidemiology Training Program entrance examination. Model responses were evaluated against standard answers and benchmarked against performance scores from junior epidemiologists.

Results For multiple-choice questions, only DeepSeek-V3 (75%) exceeded the 75th percentile performance level of junior epidemiologists (67.5%). In case-based assessments, most LLMs achieved or surpassed the 75th percentile of junior epidemiologists, demonstrating particular strength in data analysis tasks.

Conclusion Although LLMs demonstrate promise as supportive tools in field epidemiology investigations, they cannot yet replace human expertise. Significant challenges persist regarding the accuracy and timeliness of model outputs, alongside critical concerns about data security and privacy protection that must be addressed before widespread implementation.

Field epidemiology investigation serves as a cornerstone of public health practice, proving essential for identifying risk factors and implementing effective control measures. Large language models (LLMs) have recently emerged as potentially transformative tools in this domain (*1*). Models such as ChatGPT and DeepSeek have demonstrated impressive capabilities in text generation, reasoning, and data analysis. These systems can interpret user commands and generate contextually appropriate responses, positioning LLMs as valuable support tools across diverse fields.

Previous research has primarily concentrated on clinical applications of LLMs, where they have shown promise in medical diagnosis, patient counseling, and medical record management (*2*). While these applications highlight the broad potential of LLMs, their effectiveness in supporting field epidemiology investigations remains uncertain. Field epidemiology investigation encompasses extensive knowledge domains, including clinical medicine, epidemiology, laboratory and behavioral sciences, laws and regulations, technical guidelines, and decision-making frameworks (*3*). The existing literature on LLMs in public health remains limited, with few studies specifically examining their role in field epidemiology investigations. Moreover, most research has been conducted in Western contexts, leaving the application of LLMs in field epidemiology investigations — particularly within Chinese-language environments — largely unexplored. Given the rapid advancement of artificial intelligence (AI) Plus initiatives, investigating how LLMs can assist epidemiological investigations carries significant practical importance.

This study addressed this knowledge gap by evaluating the performance of several leading LLMs in executing common field epidemiology investigation tasks. The research not only contributes to a broader understanding of LLM applications in public health but also provides valuable insights for developing AI-assisted tools for field epidemiology investigations in China.

We selected six leading large language models for evaluation: three reasoning models (ChatGPT-o4-mini-high, DeepSeek-R1, and Qwen3-235B-A22B) and three non-reasoning models (DeepSeek-V3, Qwen2.5-max, and ChatGPT-4o). ChatGPT-o4-mini-high and ChatGPT-4o are proprietary closed-source models, while the remaining four represent open-source alternatives. Our evaluation framework utilized questions from the 2025 Zhejiang Field Epidemiology Training Program entrance examination, with all materials reviewed by field epidemiology experts to ensure accuracy and clarity. A total of 35 junior field epidemiologists participated in the examination. The assessment comprised two components: multiple-choice questions testing foundational knowledge and case-based scenarios evaluating practical application skills. The multiple-choice section included 20 single-answer questions with five options each, covering core topics such as infectious disease surveillance and reporting, risk assessment, outbreak management protocols, and sample collection procedures. The case-based questions presented open-ended scenarios requiring sequential responses, with each subsequent question posed only after the model completed its previous answer. This approach simulates real-world outbreak response conditions and evaluates the models' capacity to provide accurate, professional guidance on demand. All models were accessed on May 12, 2025, through their respective web interfaces using standardized Chinese-language prompts. Additional methodological details are available in the Supplementary Material (available at https://weekly.chinacdc.cn/).

For multiple-choice questions, we compared model responses against standard answers, awarding one point for each correct response (maximum score: 20 points). The case-based section contained four questions, with each response independently evaluated by two expert assessors. These evaluators scored responses against established criteria, including scientific accuracy, comprehensiveness, clarity of presentation, and contextual relevance. Each open-ended question carried a maximum score of 10 points.

For the multiple-choice questions, we calculated the proportion of correct answers for each LLM and compared these results with responses from junior epidemiologists. Statistical differences were assessed using binomial tests with p_0=0.20 (LLMs versus chance) and bootstrap approaches (highest-scoring LLM versus junior epidemiologists). For the case-based questions, we computed Pearson's *r* and Spearman's ρ to evaluate the correlation between the two evaluators' ratings. We conducted Friedman and Wilcoxon tests to examine score differences in the open-ended questions. All statistical analyses were performed using the "stats" package in R software (version 4.3.2, R Core Team, Vienna, Austria). Statistical significance was set at $P\leq0.05$.

Figure 1 demonstrates the performance of each LLM on the multiple-choice questions. Among the 20 questions, DeepSeek-V3 and Qwen3-235B-A22B achieved the highest scores, with 15/20 [75%, 95% confidence interval (*CI*): 50.9%, 91.3%] and 13/20 (65%, 95% *CI*: 40.8%, 84.6%), respectively. ChatGPT-o4-mini-high and ChatGPT-4o obtained the lowest scores, both scoring 8/20 (40%, 95% *CI*: 19.1%, 63.9%). The results revealed that four models achieved accuracy rates significantly higher than random guessing (P<0.05), with the exceptions being ChatGPT-o4-mini-high and ChatGPT-4o. Additional results are provided in the supplementary materials (Figure S1 and Table S1). When comparing the top-performing model, DeepSeek-V3 demonstrated significantly better performance than the median

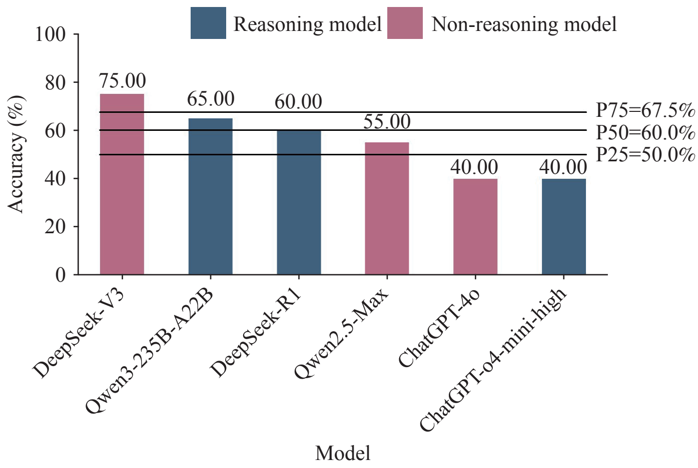


FIGURE 1. Accuracy of each model in multiple-choice questions.

accuracy rate of junior epidemiologists (60.0%) (P<0.05).

Table 1 demonstrates strong inter-rater reliability between the two experts in their evaluation of the case-based questions. Consequently, we utilized the average of both evaluators' scores as the final assessment for each open-ended question.

In the case-based section, performance varied across questions and models. For Question 1, DeepSeek-V3 achieved the highest score and was the only model to exceed the 75th percentile (P75) of junior epidemiologist scores. For Question 2, ChatGPT-4o demonstrated superior performance, while Qwen2.5 and DeepSeek-R1 both matched the P75 level of junior epidemiologists. For Question 3, four models — Qwen2.5, Qwen3-235B-A22B, DeepSeek-V3, and DeepSeek-R1 — all scored above the P75 level of junior epidemiologists, with Qwen2.5 achieving the highest score. For Question 4, all models except ChatGPT-o4-mini-high exceeded the P75 level of junior epidemiologists, with ChatGPT-4o demonstrating the strongest performance.

The chi-squared value from the Friedman test was 6.765, with a P of 0.239. Paired Wilcoxon tests revealed that the differences between ChatGPT-o4-mini-high and the other five models (DeepSeek-R1, P=0.11; DeepSeek-V3, P=0.11; Qwen3-235B-A22B, P=0.11, Qwen2.5-Max, P=0.10, ChatGPT-4o, P=0.34) were not statistically significant. All other pairwise comparisons yielded P greater than 0.5.

DISCUSSION

This study evaluated the capabilities of six currently popular LLMs in supporting field epidemiology investigations and compared their performance with examination scores from junior field epidemiologists. Among the multiple-choice questions, DeepSeek-V3 achieved the highest accuracy rate, followed by Qwen3-235B-A22B and DeepSeek-R1. For the case-based questions, no statistically significant differences were observed among the models overall; however, ChatGPT-o4-mini-high demonstrated relatively poor performance compared to the other models.

In this study, the Chinese-language LLMs (DeepSeek and Qwen) demonstrated superior performance compared to ChatGPT. The DeepSeek and Qwen models were developed using extensive Chinese language corpora during training, whereas ChatGPT was trained with limited Chinese-language content (*4*). Consequently, ChatGPT performed poorly on questions that relied heavily on Chinese language knowledge or cultural context. However, for tasks such as data analysis (Question 4), which are less dependent on Chinese-language training data, ChatGPT exhibited acceptable performance.

This study revealed that, for multiple-choice questions, most LLMs achieved lower accuracy rates than the 75th percentile level of junior field epidemiologists. Conversely, in the case-based

TABLE 1. Correlation between scores assigned by two evaluators for responses provided by six large language models.

Question	Pearson correlation		Spearman correlation	
	r	P	ρ	P
Question 1	0.937	0.006	0.742	0.091
Question 2	0.859	0.028	0.857	0.029
Question 3	0.860	0.028	0.739	0.094
Question 4	0.970	0.001	0.953	0.003

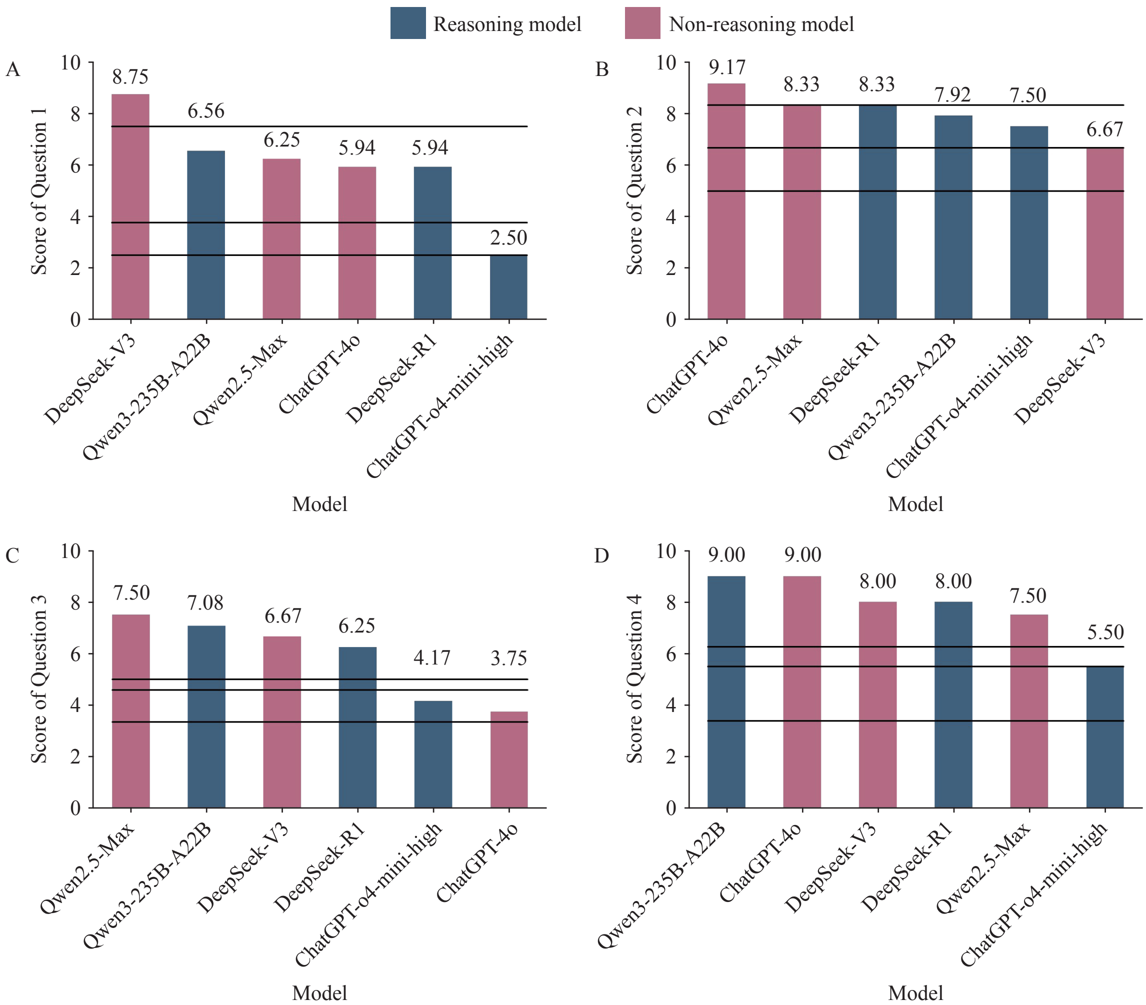


FIGURE 2. The average scores for the answers to each open-ended question provided by the six large language models. (A) Average Scores for Question 1; (B) Average Scores for Question 2; (C) Average Scores for Question 3; (D) Average Scores for Question 4.

questions, the overall performance of LLMs exceeded that of most junior field epidemiologists. Poor performance was particularly evident on Question 1, which involved professional prevention and control protocols for specific infectious diseases. This limitation may primarily stem from the absence of specialized knowledge resources in the LLMs (*4*). Similarly, the LLMs scored relatively low on Question 3, which addressed outbreak control measures. Their responses frequently included irrelevant or non-essential content, likely due to the same knowledge gap, resulting in answers that lacked technical precision and professional rigor.

Previous research has indicated that closed-source models may outperform their open-source counterparts (*5*). However, our findings demonstrate that the four Chinese open-source models generally exceeded ChatGPT's performance, underscoring the substantial potential of open-source architectures. Open-source models offer the advantage of local deployment, providing enhanced data security — a feature of paramount importance for developing specialized LLMs tailored to public health institutions.

Our study also revealed that reasoning models did not demonstrate superior performance compared to non-reasoning models, a finding consistent with observations by Sandmann et al. (*6*). Through chain-of-thought prompting in the reasoning models, we observed that LLMs incorporate knowledge from various temporal periods within their training datasets. However, these models lack the capability to distinguish between outdated and current information, resulting in instances where they failed to provide the most up-to-date knowledge.

Nevertheless, the implementation of LLMs in field epidemiology investigations continues to face several significant challenges. A critical concern is that field epidemiology is intrinsically linked to disease prevention and control, which demands exceptional timeliness and accuracy in model outputs. Our investigation identified limitations regarding citation accuracy in LLM-generated responses. In the case-based questions, several LLMs referenced guidelines or technical documents that were entirely fabricated. This presents substantial risks for junior professionals who may depend on these models without possessing the expertise to identify such erroneous references. Furthermore, LLMs trained on public knowledge bases carry an inherent risk of data contamination, potentially compromising the reliability of their outputs. These limitations have been documented in the existing literature (*7*–*8*). We therefore strongly recommend that professionals exercise caution when utilizing LLMs, cross-reference their outputs against established trusted sources, and treat these models as supplementary tools rather than substitutes for individual knowledge and experience. To enhance model performance, developing specialized knowledge resources for LLMs will be essential, supported by high-quality, regularly updated datasets for training purposes.

Another critical challenge involves data security and privacy protection (*9*). Field epidemiology investigations frequently handle sensitive information, including patient privacy data and confidential government decision-making processes, all requiring robust protection measures. Without adequate safeguards, the practical implementation of LLMs could face severe limitations. To address these concerns, comprehensive regulatory frameworks will play an essential role. The European Union has already established relevant regulations through the *EU AI Act*, representing the world's first comprehensive artificial intelligence legislation. In 2023, China also issued *China's Interim Measures for the Management of Generative AI Services*. However, as an emerging technology, LLM governance and oversight require continued research and development to ensure both innovation advancement and safety assurance (*10*).

This study presents several limitations that warrant consideration. First, our evaluation was restricted to entrance exam questions from the Zhejiang Field Epidemiology Training Program, which may not comprehensively represent all aspects of field epidemiology investigations. Second, LLM outputs exhibit inherent stochasticity, meaning responses to identical prompts may vary across individual runs. However, existing research indicates that for knowledge-intensive tasks, while model performance may show sensitivity to minor prompt variations, it generally maintains relative stability overall. Finally, our evaluation employed a limited number of questions, with case-based scenarios focusing exclusively on infectious diseases. Consequently, model performance in other types of public health emergencies remains uncertain. Future studies should expand the evaluation scope to enhance the reliability and generalizability of these findings.

This study evaluated the potential of six leading LLMs to support field epidemiology investigations by comparing their performance against junior field epidemiologists' examination scores. Our findings demonstrate that several models achieved notable accuracy and relevance across both multiple-choice and case-based assessments. However, current LLMs cannot yet replace human epidemiological expertise. While these models show promise as supplementary tools, their practical implementation faces significant challenges. Future development should prioritize integrating verified knowledge databases to optimize model performance and establishing robust regulatory frameworks to ensure their safe and effective application in public health settings.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the Medical Health Science and Technology Project of Zhejiang Provincial Health Commission (2024KY896), the Public Health Talent Development Support Program of National Disease Control and Prevention Administration (2025) and the Zhejiang Provincial Key Laboratory of Vaccine and Preventive Research on Infectious Diseases from Department of Science and Technology of Zhejiang Province (2024).

doi: 10.46234/ccdcw2025.220

Corresponding author: Chen Wu, chenwu@cdc.zj.cn.

[1] Department of Public Health Surveillance and Advisory, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou City, Zhejiang Province, China; [2] Department of Communicable Disease Control and Prevention, Shaoxing Center for Disease Control and Prevention, Shaoxing City, Zhejiang Province, China; [3] Department of Communicable Disease Control and Prevention, Zhoushan Center for Disease Control and Prevention, Zhoushan City, Zhejiang Province, China.

& Joint first authors.

Submitted: June 20, 2025
Accepted: September 27, 2025
Issued: October 10, 2025

REFERENCES

1. Conroy G, Mallapaty S. How China created AI model DeepSeek and shocked the world. Nature 2025;638(8050):300 – 1. https://doi.org/10.1038/d41586-025-00259-0.
2. Yim D, Khuntia J, Parameswaran V, Meyers A. Preliminary evidence of the use of generative AI in health care clinical services: systematic narrative review. JMIR Med Inform 2024;12:e52073. https://doi.org/10.2196/52073.
3. Rasmussen SA, Goodman RA. The CDC field epidemiology manual. New York: Oxford University Press. 2018. https://www.amazon.com/CDC-Field-Epidemiology-Manual/dp/0190624248.
4. Wu JG, Wu X, Qiu ZP, Li MH, Lin SX, Zhang YY, et al. Large language models leverage external knowledge to extend clinical insight beyond language boundaries. J Am Med Inform Assoc 2024;31(9):2054 – 64. https://doi.org/10.1093/jamia/ocae079.
5. Nazi ZA, Hossain R, Mamun FA. Evaluation of open and closed-source LLMs for low-resource language with zero-shot, few-shot, and chain-of-thought prompting. Natl Lang Process J 2025;10:100124. https://doi.org/10.1016/j.nlp.2024.100124.
6. Sandmann S, Hegselmann S, Fujarski M, Bickmann L, Wild B, Eils R, et al. Benchmark evaluation of DeepSeek large language models in clinical decision-making. Nat Med 2025;31(8):2546 – 9. https://doi.org/10.1038/s41591-025-03727-2.
7. Clelland CL, Moss S, Clelland JD. Warning: artificial intelligence chatbots can generate inaccurate medical and scientific information and references. Explor Digit Health Technol 2024;2:1 – 6. https://doi.org/10.37349/edht.2024.00006.
8. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare 2023;11(6):887. https://doi.org/10.3390/healthcare11060887.
9. The Lancet Digital Health. ChatGPT: friend or foe? Lancet Digit Health 2023;5(3):e102. http://dx.doi.org/10.1016/s2589-7500(23)00023-7.
10. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. npj Digit Med 2023;6(1):120. https://doi.org/10.1038/s41746-023-00873-0.

SUPPLEMENTARY MATERIALS

Access Date and Settings for the Large Language Models

All LLMs were accessed through their web interfaces on May 12, 2025. The DeepSeek-V3 model was the 0324 build version, whereas specific build versions were not disclosed for DeepSeek-R1 (initial release), Qwen, or ChatGPT. No additional tools or plugins were used. All reasoning models displayed the chain of thought by default. Each model was queried using a newly registered account that had not been used for any prior interactions, ensuring no influence from historical usage or personalization. The prompts did not include chain-of-thought or “reasoning” related instructions. To ensure fairness, all chat memory and user personalization settings were disabled. This prevented models from benefiting from prior context and guaranteed that each query was processed independently. The specific settings were as follows:

Qwen: Switched to the test model with the temporary conversation setting enabled. Internet search was disabled; all other settings remained at their defaults.

DeepSeek: Switched to the test model. Internet search was disabled, with all other settings kept as defaults.

ChatGPT: Switched to the test model with the temporary conversation setting enabled. All other settings kept as defaults.

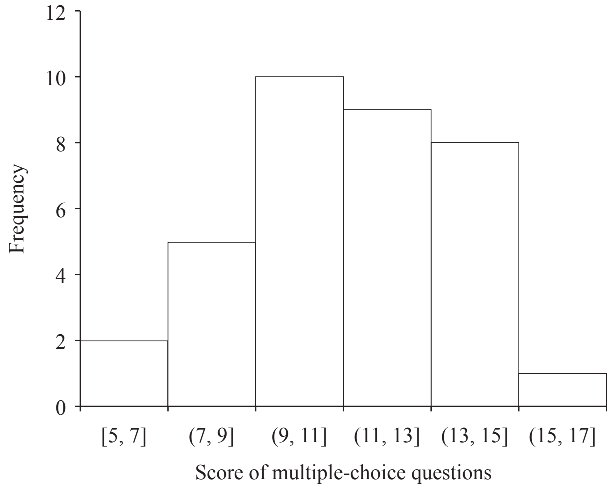
The prompts were as follows:

For case-based questions: “Please play the role of an on-site epidemiological investigation expert from the CDC and answer the following question”.

For multiple-choice questions: “Below is the written examination for the enrollment of a field epidemiology training program at a provincial CDC. Please provide your answers and indicate which ones you believe to be correct”.

Performance of the Junior Epidemiologists on the Multiple-Choice Questions

A total of 35 junior epidemiologists participated in the examination, with a mean score of 11.4 (standard deviation =2.6). The score distribution was as follows: minimum =5/20, 25th percentile (P25) =10/20, median (P50) =12/20, 75th percentile (P75) =13.5/20, and maximum =16/20.



SUPPLEMENTARY FIGURE S1. Histogram of score distribution for 35 junior epidemiologists who took the examination.

SUPPLEMENTARY TABLE S1. Performance of the six large language models on the multiple-choice questions.

Models	**Accuracy (%)**	**95% confidence interval (%)**	***P****
DeepSeek-V3	75.0	50.9, 91.3	<0.001
Qwen-235B-A22B	65.0	40.8, 84.6	<0.001
DeepSeek-R1	60.0	36.1, 80.9	<0.001
Qwen2.5-Max	55.0	31.5, 76.9	<0.001
ChatGPT-4o	40.0	19.1, 63.9	0.262
ChatGPT-o4-mini-high	40.0	19.1, 63.9	0.262

* Two-sided Bonferroni-adjusted *P* for comparing the model with chance.

Preplanned Studies

Spatial Distribution and Clustering Patterns of Cognitive Impairment Among the Older Population — 31 PLADs, China, 2024

Ying Liu[1,&]; Yushan Zhang[2,&]; Ji Shen[2]; Chi Zhang[3]; Yingchen Sang[1]; Youjiao Wang[1]; Houguang Zhou[4]; Lei Wang[5]; Jie Zhang[2]; Ying Yuan[2]; Shiwei Liu[1,#]; Hong Shi[2,#]

Summary

What is already known about this topic?

As the Chinese population ages, the prevalence of cognitive impairment among older adults continues to increase. Cognitive impairment severely restricts daily activities and creates significant social and economic burdens.

What is added by this report?

Using nationally representative data from the China Survey of Aging and Health (CAHS), this study found that the weighted prevalence of subjective cognitive decline and mild cognitive impairment among individuals aged 65 years and older in China was 38.8% and 28.4% in 2024, respectively, and both showed spatial clustering.

What are the implications for public health practice?

Through the analysis of spatial distribution patterns and identification of high-risk regions of cognitive impairment, this study provides critical information for developing targeted regional prevention and control interventions.

ABSTRACT

Introduction: Cognitive impairment poses a serious threat to the health of older adults. Understanding spatial distribution patterns and identifying high-risk areas are essential for developing targeted regional prevention and control strategies. This study examined the spatial distribution and clustering patterns of cognitive impairment in China in 2024.

Methods: This study utilized data from the 2024 China Survey of Aging and Health. Rao-Scott chi-square tests were used to compare differences in prevalence across demographic subgroups. Global and local spatial autocorrelation analyses were conducted to examine the spatial distribution patterns and clustering characteristics.

Results: In 2024, the prevalence of subjective cognitive decline (SCD) and mild cognitive impairment (MCI) among older adults ≥65 years in China was 38.8% and 28.4%, respectively. The prevalence of SCD was highest in western China (45.1%), while MCI was highest in central China (31.0%). Global spatial autocorrelation analysis revealed that SCD (P=0.025) and MCI (P=0.015) distribution exhibited spatial clustering across China.

Conclusions: The current burden of cognitive impairment in China's older population is substantial and characterized by significant regional variations. Prevention and treatment measures should prioritize support for high-prevalence areas with limited resources and promote scientifically based, precise, and efficient cognitive impairment prevention and treatment strategies throughout China.

Cognitive impairment is a cardinal manifestation of neurodegenerative disorders, including Alzheimer's disease (AD) and other dementia syndromes (*1*). Subjective cognitive decline (SCD) is a preclinical stage of AD, with progressive SCD potentially advancing to mild cognitive impairment (MCI), which carries an elevated risk of further progression to AD (*2*). Current research indicates that approximately 15.07 million individuals >60 years in China live with dementia (*3*). An epidemiological study conducted from 2018 to 2019 documented an MCI prevalence rate of 27.8% among individuals ≥65 years (*4*). This study utilized data from the China Survey of Aging and Health (CAHS) to examine the prevalence of cognitive impairment among adults aged ≥65 years throughout China. Through a systematic analysis of spatial distribution patterns and identification of high-risk geographic regions, this study aimed to inform the development of targeted regional prevention and intervention strategies.

The CAHS utilized a multistage stratified cluster sampling design across all 31 provincial-level administrative divisions (PLADs) in China. This study utilized probability proportional-to-size sampling to ensure both national and provincial representativeness. The sampling framework comprises three sequential stages. First, the required sample size for each PLAD was calculated based on the standardized age distribution and sex ratio of the older population. Second, two to three representative survey cities were selected from each PLAD based on prefecture-level city gross domestic product rankings (with municipalities required to include two urban and two suburban districts); each prefecture-level city contributed a minimum of two districts or counties representing medium development levels, with each district providing two communities for survey participation. The CAHS successfully recruited 49,193 individuals aged ≥65 years. After implementing rigorous onsite quality control measures and comprehensive data cleaning procedures, 41,859 valid questionnaires were included in the final analysis, yielding an overall response rate of 85.1%. The Beijing Hospital Ethics Committee approved the study protocol (approval number: 2021BJYEC-114-01), and informed consent was obtained from all participants.

The CAHS utilized the Subjective Cognitive Decline Nine-item Questionnaire to assess patients' subjective experience of cognitive decline, despite objective test performance remaining within normal ranges for their age and education level, without meeting the criteria for MCI or dementia (*5*). Additionally, the Ascertain Dementia Eight-item Questionnaire demonstrates high sensitivity and specificity in identifying initial cognitive changes associated with various types of dementia, including AD, vascular dementia, dementia with Lewy bodies, and frontotemporal dementia, while remaining unaffected by age, education, or cultural differences (*6*).

Statistical analyses included sampling design, non-response adjustment, and post-stratification correction weights to ensure population representativeness. We computed prevalence estimates with 95% confidence intervals (*CI*s) using survey-weighted methods and evaluated subgroup disparities through Rao-Scott χ^2 tests. All descriptive analyses were conducted using STATA (version 18.0; StataCorp LLC, College Station, Texas, USA). We used ArcGIS (version 10.8.1; Environmental Systems Research Institute, California, USA) to conduct spatial distribution and clustering pattern analyses. The spatial weights matrix was defined using the "Inverse Distance" method for conceptualizing spatial relationships, with the distance method set to the default "Euclidean_Distance" and the standardization method set to "ROW." Using PLADs as spatial analysis units, we employed global spatial autocorrelation analysis to assess the overall spatial aggregation of cognitive impairment across the 31 PLADs. We used local spatial autocorrelation analyses to identify clustering patterns, categorizing the regions into four distinct types: high-high, low-high, low-low, and high-low aggregation areas. Statistical significance was set at P<0.05.

In 2024, the weighted prevalence of SCD and MCI among older adults aged ≥65 years in China was 38.8% and 28.4%, respectively. SCD prevalence was highest in western China (45.1%), whereas that of MCI peaked in central China (31.0%). Females demonstrated higher prevalence rates for both conditions (39.7% and 28.4%, respectively) than males (37.8% and 28.3%, respectively). Rural areas consistently exceeded urban areas in terms of the prevalence rates (41.7% *vs.* 37.1% for SCD and 33.8% *vs.* 25.1% for MCI). Adults aged ≥80 years exhibited the highest prevalence rates for both SCD (46.3%) and MCI (40.3%). Higher educational attainment demonstrated a significant inverse association between SCD prevalence and MCI incidence rates. When body mass index fell within the normal range of 18.5–24.9 kg/m^2, the prevalence of SCD was lowest (36.9%) (Table 1).

In 2024, SCD prevalence among adults aged ≥65 years across China's 31 PLADs ranged from 23.6% (Hainan) to 72.4% (Qinghai). Global spatial autocorrelation analysis revealed that the distribution of SCD exhibits spatial clustering across China (Moran's I=0.162, Z=2.242, P=0.025). Local spatial autocorrelation analysis identified distinct clustering patterns: high-high clustering areas: Ningxia, Xizang, Sichuan, and Gansu; low-low clustering areas: Liaoning and Hebei; high-low clustering area: Heilongjiang. The geographic distribution revealed that SCD among older adults is predominantly concentrated in China's western regions (Table 2).

In 2024, MCI prevalence among individuals aged ≥65 years across China's 31 PLADs ranged from 12.3% (Tianjin) to 58.9% (Qinghai). Global spatial autocorrelation analysis revealed that the distribution of MCI exhibits spatial clustering across China (Moran's I=−0.242, Z=−2.431, P=0.015). Local spatial

TABLE 1. Demographic distribution of subjective cognitive decline and mild cognitive impairment among adults aged 65 years and older population in 31 PLADs, China, 2024.

Characteristics	Participants *N* (%)	Subjective Cognitive Decline		Mild Cognitive Impairment	
		Weighted % (95% *CI*)	*P*	Weighted % (95% *CI*)	*P*
Total	41,859 (100.0)	38.8 (35.7, 42.1)		28.4 (25.3, 31.6)	
Region			<0.001		<0.001
Eastern	16,795 (40.1)	33.5 (28.5, 39.0)		27.7 (21.7, 34.5)	
Central	11,462 (27.4)	41.5 (35.7, 47.6)		31.0 (26.3, 36.2)	
Western	10,115 (24.2)	45.1 (39.0, 51.3)		28.4 (23.8, 33.6)	
Northeastern	3,487 (8.3)	37.1 (29.1, 45.9)		24.6 (18.9, 31.4)	
Sex			<0.001		<0.001
Male	19,893 (47.5)	37.8 (33.4, 42.4)		28.3 (23.6, 33.6)	
Female	21,966 (52.5)	39.7 (35.6, 44.0)		28.4 (24.7, 32.4)	
Age (years)			<0.001		<0.001
65–69	12,518 (29.9)	33.9 (28.7, 39.6)		23.8 (19.2, 29.1)	
70–74	12,794 (30.6)	35.7 (30.2, 41.6)		23.5 (19.2, 28.4)	
75–79	8,001 (19.1)	44.3 (38.2, 50.6)		30.3 (25.0, 36.3)	
≥80	8,546 (20.4)	46.3 (39.0, 53.8)		40.3 (32.3, 48.9)	
Area type			<0.001		<0.001
Urban	25,553 (61.1)	37.1 (33.1, 41.1)		25.1 (21.8, 28.6)	
Rural	16,306 (38.9)	41.7 (36.4, 47.2)		33.8 (28.0, 40.0)	
Education			<0.001		<0.001
Primary school or less	22,475 (53.7)	42.7 (37.9, 47.6)		34.1 (29.1, 39.4)	
Secondary school	10,402 (24.8)	35.7 (30.6, 41.2)		22.7 (18.8, 27.3)	
High school and above	8,982 (21.5)	33.9 (28.6, 39.6)		21.9 (17.4, 27.2)	
BMI (kg/m^2)			0.004		0.065
<18.5	1,920 (4.6)	42.8 (30.8, 55.8)		34.7 (23.9, 47.4)	
18.5–24.9	27,769 (66.3)	36.9 (33.2, 40.7)		27.3 (23.6, 31.4)	
25.0–29.9	10,765 (25.7)	42.0 (35.6, 48.6)		29.3 (23.7, 35.6)	
≥30.0	1,405 (3.4)	45.2 (30.0, 61.3)		32.2 (19.6, 48.1)	
Monthly household income (CNY)			<0.001		<0.001
<3,000	11,437 (27.3)	42.3 (35.9, 49.0)		30.3 (25.0, 36.3)	
3,000–5,999	11,468 (27.4)	38.4 (32.9, 44.1)		26.3 (21.0, 32.4)	
6,000–9,999	7,989 (19.1)	37.2 (31.8, 42.9)		25.0 (20.2, 30.5)	
≥10,000	4,703 (11.2)	33.6 (27.1, 40.8)		25.8 (20.0, 32.6)	
Unwilling to disclose	6,262 (15.0)	41.8 (31.1, 53.4)		37.4 (24.8, 51.9)	

Abbreviation: PLADs=provincial-level administrative divisions; BMI=body mass index; CNY=Chinese Yuan; *CI*=confidence interval.

autocorrelation analysis identified Yunnan as a high-low clustering area and Hebei as a low-high clustering area (Table 3).

DISCUSSION

This study found that the prevalence of SCD and MCI among older adults ≥65 years in China in 2024 was 38.8% and 28.4%, respectively. Previous domestic research revealed an MCI prevalence rate of 27.8% among the same age group (*4*). The 28.4% prevalence of MCI among China's older population in 2024 found in this study indicates that the prevention and treatment of cognitive impairment represents a substantial public health challenge that requires urgent attention. Additionally, SCD prevalence was highest in

TABLE 2. Prevalence and clustering patterns of subjective cognitive decline among adults aged 65 years and older across 31 PLADs, China, 2024.

PLADs	Sample size, *N* (%)	Weighted prevalence (%)	*Z* Score	*P*	Clustering patterns
Hainan	199 (0.5)	23.6 (7.4, 54.3)	0.968	0.176	
Guangdong	2,063 (4.9)	24.9 (14.2, 40.0)	1.052	0.160	
Beijing	1,216 (2.9)	26.1 (19.6, 34.0)	0.844	0.208	
Tianjin	453 (1.1)	26.4 (17.4, 38.0)	1.011	0.164	
Liaoning	1,639 (3.9)	27.0 (12.8, 48.4)	1.633	0.050	LL
Zhejiang	1,775 (4.2)	28.1 (14.2, 47.8)	−0.052	0.500	
Jilin	802 (1.9)	29.0 (20.4, 39.5)	1.264	0.098	
Guangxi	1,372 (3.3)	29.8 (17.6, 45.7)	0.479	0.342	
Henan	3,342 (8.0)	35.7 (26.4, 46.1)	−0.852	0.212	
Jiangxi	1,031 (2.4)	36.7 (26.3, 48.7)	0.199	0.418	
Chongqing	1,124 (2.7)	36.9 (21.7, 55.1)	−1.625	0.058	
Inner Mongolia	661 (1.6)	38.6 (17.6, 65.0)	1.331	0.102	
Fujian	899 (2.1)	38.7 (24.2, 55.5)	1.203	0.118	
Hebei	2,394 (5.7)	38.7 (29.3, 49.1)	1.796	0.018	LL
Jiangsu	2,746 (6.6)	39.8 (26.4, 55.0)	0.221	0.418	
Hunan	2,619 (6.3)	40.2 (24.6, 58.0)	0.271	0.396	
Ningxia	150 (0.4)	40.7 (24.2, 59.6)	2.118	0.014	HH
Yunnan	1,027 (2.4)	41.5 (17.2, 70.9)	0.304	0.376	
Shanxi	903 (2.2)	42.6 (26.2, 60.8)	−0.016	0.498	
Xizang	41 (0.1)	43.1 (25.2, 63.0)	2.818	0.006	HH
Sichuan	2,672 (6.4)	43.2 (28.8, 58.9)	2.290	0.016	HH
Shandong	3,104 (7.4)	44.2 (28.1, 61.7)	−1.252	0.108	
Shanghai	1,946 (4.6)	45.1 (33.6, 57.1)	−0.909	0.192	
Anhui	1,805 (4.3)	45.6 (37.5, 53.9)	−0.666	0.254	
Heilongjiang	1,046 (2.5)	48.2 (37.6, 58.9)	−1.578	0.050	HL
Xinjiang	408 (1.0)	48.9 (40.7, 57.3)	0.256	0.370	
Hubei	1,762 (4.2)	49.8 (32.3, 67.4)	−0.240	0.404	
Gansu	624 (1.5)	51.5 (27.9, 74.5)	2.754	0.010	HH
Guizhou	866 (2.1)	55.3 (40.3, 69,3)	−0.546	0.298	
Shaanxi	1,046 (2.5)	62.6 (53.0, 71.3)	0.580	0.294	
Qinghai	124 (0.3)	72.4 (55.1, 84.9)	1.307	0.094	

Abbreviation: HH=high-high; LH=low-high; LL=low-low; HL=high-low; PLADs=provincial-level administrative divisions.

the western regions, whereas MCI prevalence peaked in the central regions, consistent with previous research findings. These significant regional disparities in prevalence likely reflect regional differences in economic development, availability of medical resources, and access to health education.

We observed significant regional differences in SCD and MCI prevalence among adults ≥65 years across 31 provinces in China. Notably, Qinghai Province exhibited the highest prevalence of both SCD and MCI compared to Xizang. Both Qinghai and Xizang are located on the Tibetan Plateau, where altitude gradients (1,500–2,500 m, 2,500–4,000 m, and ≥ 4,000 m) demonstrate significant associations with cognitive function changes. These effects are moderated by residential history (long-term/lifetime residence) and acclimatization levels, suggesting that the impact of high-altitude hypoxia on cognition may vary across individuals and environments (*7*). The hypoxic conditions in Qinghai are generally less severe than those in Xizang. Moderate hypoxia may pose greater risks than extreme hypoxia: chronic "sub-lethal"

TABLE 3. Prevalence and clustering patterns of mild cognitive impairment among adults aged 65 years and older across 31 PLADs, China, 2024.

PLADs	Sample size, *N* (%)	Weighted prevalence (%)	*Z* Score	*P*	Clustering patterns
Tianjin	453 (1.1)	12.3 (4.4, 29.8)	−0.845	0.214	
Hainan	199 (0.5)	16.1 (4.6, 43.1)	0.120	0.418	
Xizang	41 (0.1)	17.2 (8.6, 31.5)	0.637	0.290	
Ningxia	150 (0.4)	19.4 (14.4, 25.5)	0.357	0.410	
Jiangsu	2,746 (6.6)	19.4 (11.3, 31.3)	1.114	0.134	
Liaoning	1,639 (3.9)	20.1 (9.2, 38.4)	−0.566	0.264	
Jilin	802 (1.9)	20.8 (14.7, 28.6)	0.973	0.170	
Beijing	1,216 (2.9)	20.8 (12.4, 32.8)	0.264	0.384	
Zhejiang	1,775 (4.2)	21.1 (9.6, 40.2)	−0.817	0.210	
Inner Mongolia	661 (1.6)	21.9 (8.6, 45.7)	0.751	0.214	
Sichuan	2,672 (6.4)	22.5 (12.7, 36.6)	−0.990	0.170	
Chongqing	1,124 (2.7)	23.5 (11.7, 41.6)	0.757	0.238	
Henan	3,342 (8.0)	24.2 (17.2, 32.7)	−0.130	0.418	
Yunnan	1,027 (2.4)	25.2 (10.0, 50.5)	−1.429	0.050	HL
Gansu	624 (1.5)	26.4 (14.3, 43.5)	−0.968	0.164	
Fujian	899 (2.1)	27.3 (13.8, 46.7)	0.609	0.254	
Hubei	1,762 (4.2)	28.4 (15.4, 46.4)	−0.953	0.170	
Hunan	2,619 (6.3)	28.7 (18.2, 42.0)	−0.192	0.384	
Guangxi	1,372 (3.3)	28.9 (15.9, 46.8)	−0.755	0.248	
Hebei	2,394 (5.7)	29.2 (22.3, 37.1)	−3.135	0.004	LH
Jiangxi	1,031 (2.4)	29.6 (20.3, 41.0)	−0.946	0.186	
Shanghai	1,946 (4.6)	29.6 (17.3, 45.8)	−1.225	0.098	
Heilongjiang	1,046 (2.5)	29.7 (21.0, 40.3)	−0.265	0.426	
Guangdong	2,063 (4.9)	34.6 (17.3, 57.2)	−0.261	0.426	
Shaanxi	1,046 (2.5)	35.9 (28.2, 44.3)	−1.120	0.120	
Xinjiang	408 (1.0)	36.6 (29.4, 44.4)	−1.134	0.060	
Guizhou	866 (2.1)	38.9 (26.6, 52.7)	0.466	0.330	
Shanxi	903 (2.2)	40.7 (25.3, 58.3)	−1.804	0.054	
Anhui	1,805 (4.3)	41.2 (31.1, 52.2)	−0.345	0.418	
Shandong	3,104 (7.4)	42.7 (25.6, 61.7)	−0.186	0.450	
Qinghai	124 (0.3)	58.9 (38.7, 76.4)	−1.254	0.102	

Abbreviation: HL=high-low; LH=low-high; PLADs=provincial-level administrative divisions.

hypoxia can lead to sustained oxidative stress and neuroinflammation, whereas extreme hypoxia may trigger more robust protective physiological responses (*8*).

Global spatial autocorrelation analysis revealed that the distribution of SCD across China exhibited clustering patterns in which high-prevalence regions were geographically adjacent to other high-prevalence areas. In contrast, MCI demonstrated a distinct spatial pattern characterized by low-prevalence regions in neighboring high-prevalence areas. This divergence between SCD and MCI stems from a fundamental misalignment between subjective self-perception and objective cognitive assessment results. This discrepancy is primarily attributable to the substantial heterogeneity within SCD populations, where symptom reporting rates depend heavily on individual awareness of cognitive changes. Consequently, individuals may report symptoms of cognitive decline that have not yet been detected (*9*). Additionally, multiple studies have demonstrated that MCI detection rates correlate strongly with healthcare accessibility, as reduced access increases the risk of underdiagnosis (*10*).

The findings in this report are subject to three limitations. First, the cross-sectional design captured cognitive status at a single time point, preventing the examination of symptom progression trajectories over time. Second, reliance on self-reported measures introduces potential recall and social desirability biases, which may compromise response accuracy. Finally, future research should incorporate specific cognitive impairment risk factors and conduct statistical analyses to examine conversion rates among SCD, MCI, and AD.

In conclusion, spatial clustering of SCD and MCI exists in the older Chinese population. Prevention and treatment strategies should emphasize regional differentiation by prioritizing support for high-burden areas with limited resources. Implementation of cognitive health literacy campaigns should improve public awareness of modifiable risk factors for cognitive impairment (e.g., hypertension, diabetes, and physical inactivity) in areas with high-high SCD clustering. Continuous monitoring of changes in SCD and MCI distributions will enable real-time evaluation of intervention effectiveness (e.g., health education and resource allocation), facilitating evidence-based adjustments and advancing precise public health initiatives.

Conflicts of interest: No conflict of interest.

Acknowledgements: All research personnel who contributed to the data collection efforts and all study participants whose involvement made this investigation possible.

Ethical statement: The CAHS protocol approved by the Ethics Committee of Beijing Hospital (approval number: 2021BJYEC-114-01).

Funding: Supported by the National Key R&D Plan "Intergovernmental International Science and Technology Innovation Cooperation" Key Special Project (2021YFE0111800).

doi: 10.46234/ccdcw2025.219

[#] Corresponding authors: Shiwei Liu, liusw@chinacdc.cn; Hong Shi, shihong2584@bjhmoh.cn.

[1] Tobacco Control Office, Chinese Center for Disease Control and Prevention, Beijing, China; [2] Department of Geriatrics, Beijing Hospital, National Center of Gerontology; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China; [3] The Key Laboratory of Geriatrics, Beijing Institute of Geriatrics, Beijing Hospital, National Center of Gerontology, National Health Commission; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China; [4] Department of Geriatrics, Huashan Hospital, National Clinical Research Center for Aging and Medicine, Fudan University, Shanghai, China; [5] Department of Geriatrics, Medical Center on Aging of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.
[&] Joint first authors.

Submitted: July 22, 2025
Accepted: September 24, 2025
Issued: October 10, 2025

REFERENCES

1. Jia LF, Quan MN, Fu Y, Zhao T, Li Y, Wei CB, et al. Dementia in China: epidemiology, clinical management, and research advances. Lancet Neurol 2020;19(1):81 – 92. https://doi.org/10.1016/S1474-4422(19)30290-X.
2. Guan H, Yue L, Yap PT, Xiao SF, Bozoki A, Liu MX. Attention-guided autoencoder for automated progression prediction of subjective cognitive decline with structural MRI. IEEE J Biomed Health Inform 2023;27(6):2980 – 9. https://doi.org/10.1109/JBHI.2023.3257081.
3. Jia LF, Du YF, Chu L, Zhang ZJ, Li FY, Lyu DY, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020;5(12):e661 – 71. https://doi.org/10.1016/S2468-2667(20)30185-7.
4. Liu D, Li L, An LN, Cheng GR, Chen C, Zou MJ, et al. Urban-rural disparities in mild cognitive impairment and its functional subtypes among community-dwelling older residents in central China. Gen Psychiatr 2021;34(5):e100564. https://doi.org/10.1136/gpsych-2021-100564.
5. Wang XQ, Huang WJ, Su L, Xing Y, Jessen F, Sun Y, et al. Neuroimaging advances regarding subjective cognitive decline in preclinical Alzheimer's disease. Mol Neurodegener 2020;15(1):55. https://doi.org/10.1186/s13024-020-00395-3.
6. Chen HH, Sun FJ, Yeh TL, Liu HE, Huang HL, Kuo BIT, et al. The diagnostic accuracy of the Ascertain Dementia 8 questionnaire for detecting cognitive impairment in primary care in the community, clinics and hospitals: a systematic review and meta-analysis. Fam Pract 2018;35(3):239 – 46. https://doi.org/10.1093/fampra/cmx098.
7. Su R, Jia SR, Zhang NN, Wang YY, Li H, Zhang DL, et al. The effects of long-term high-altitude exposure on cognition: a meta-analysis. Neurosci Biobehav Rev 2024;161:105682. https://doi.org/10.1016/j.neubiorev.2024.105682.
8. Burtscher J, Mallet RT, Burtscher M, Millet GP. Hypoxia and brain aging: neurodegeneration or neuroprotection? Ageing Res Rev 2021;68: 101343. http://dx.doi.org/10.1016/j.arr.2021.101343.
9. Nemy M, Dyrba M, Brosseron F, Buerger K, Dechent P, Dobisch L, et al. Cholinergic white matter pathways along the Alzheimer's disease continuum. Brain 2023;146(5):2075 – 88. https://doi.org/10.1093/brain/awac385.
10. Kulshreshtha A, Parker ES, Fowler NR, Summanwar D, Ben Miled Z, Owora AH, et al. Prevalence of unrecognized cognitive impairment in federally qualified health centers. JAMA Netw Open 2024;7(10): e2440411. https://doi.org/10.1001/jamanetworkopen.2024.40411.

Outbreak Reports

The Cluster of Mpox (Clade Ib) Infections — Yiwu City, Zhejiang Province, China, July–August 2025

Zhiping Long[1,&]; Liebo Zhu[2,&]; Jian Cai[3]; Hanran Ji[4]; Zhen Ye[2]; Xuguang Shi[3]; Guangming Zhang[1]; Shuying Zhu[1]; Xuelian Zhang[1]; Zhifeng Pang[1]; Kaizhi Bai[1]; Jun Jiang[2,#]; Jiming Sun[3,#]

Summary

What is already known about this topic?

Mpox transmission occurs primarily through close skin-to-skin contact, sexual contact, and exposure to contaminated materials; imported cases can establish local transmission clusters in settings characterized by high international mobility.

What is added by this report?

This report documents the first identified cluster of mpox virus Clade Ib infections in Yiwu, China, an international trade hub. A total of six laboratory-confirmed cases were identified between July 23–August 6, 2025, including one imported source case from Tanzania and five locally linked secondary cases. Among 43 environmental samples, 27 (62.8%) tested positive for MPXV DNA. Ninety close and general contacts were traced and monitored under 21-day health surveillance, with no additional infections detected. We also report a case with oropharyngeal PCR positivity prior to rash onset, suggesting potential pre-symptomatic viral shedding.

What are the implications for public health practice?

Rapid multi-agency collaboration encompassing case identification, isolation protocols, contact tracing, and environmental decontamination proves effective for interrupting transmission in international trade hubs; targeted surveillance and proactive outreach to high-risk social and sexual networks remain essential components of outbreak control.

ABSTRACT

Introduction: On August 6, 2025, the Yiwu CDC received notification of a suspected mpox case. Subsequent laboratory testing confirmed mpox virus (MPXV, Clade Ib) infection, prompting the initiation of a comprehensive multi-level epidemiologic investigation.

Methods: Between August 6–10, 2025, provincial, municipal, and county expert teams conducted systematic case finding, interview-based exposure mapping, and linkage analysis of public security and border records to reconstruct contact trajectories. The investigation included comprehensive contact tracing, environmental sampling, and real-time PCR testing of lesion and oropharyngeal specimens for MPXV detection.

Results: Six laboratory-confirmed cases were identified with symptom onset dates spanning July 23 through August 6. Five patients received treatment in Yiwu, while one was managed in Changzhou. The cases were predominantly male (5/6), with ages ranging from 22–43 years (median 30 years). Four of the six cases were foreign nationals. Investigators identified and monitored 52 core close contacts and 38 general contacts under 21-day health surveillance protocols. Environmental sampling (n=43) conducted at five case residences and personal items yielded 27 positive results (62.8%) for MPXV.

Conclusions: This outbreak represents an imported mpox cluster with subsequent person-to-person transmission occurring primarily through intimate contact. We documented substantial household environmental contamination, emphasizing the critical importance of comprehensive decontamination measures. Rapid case detection, systematic contact management, and terminal disinfection protocols effectively contained further viral spread.

Mpox is an emerging zoonotic disease caused by the mpox virus (MPXV), a member of the Orthopoxvirus genus (*1*). MPXV Clade Ib was first identified in September 2023 in the Democratic Republic of Congo and, following its rapid global dissemination, was declared a Public Health Emergency of International Concern by the WHO in August 2024 (*2*). Prior to January 2025, all reported mpox cases in China were attributed to Clade IIb; however, five MPXV Clade Ib infections were detected that month, marking the strain's initial documented entry into the country (*3*). This clade demonstrates sustained human-to-human transmission, occurring primarily through sexual contact networks (accounting for 72% of cases), household exposure, and community contact, with no evidence supporting airborne transmission (*4*). Unlike previous clades, MPXV Clade Ib exhibits enhanced virulence, with a case fatality ratio (CFR) of 5.3%. The strain disproportionately affects pediatric populations (67% of cases) and young adults in high-risk groups, including sex workers (*5*).

This report describes the first documented cluster of MPXV (Clade Ib) infections in Zhejiang Province, originating from an imported case that traveled from Tanzania to Yiwu City. We present the epidemiological, clinical, and virological characteristics of this outbreak and detail the coordinated public health interventions implemented to prevent secondary transmission.

INVESTIGATION AND RESULTS

Case Report

On August 6, 2025, the fever clinic at Yiwu Central Hospital notified Yiwu CDC of a male patient (Case 1) presenting with fever and genital rash. Considering Yiwu's substantial international population mobility and recent global mpox alerts, the attending physician suspected mpox and immediately implemented isolation protocols. Lesion and oropharyngeal swabs were collected and transported to Yiwu CDC, where same-day real-time PCR testing confirmed MPXV DNA positivity. Jinhua CDC verified these results on August 7, prompting immediate notification of the Zhejiang Provincial CDC. A joint investigation team comprising provincial, municipal, and county CDC experts was established that same day.

Epidemiologic interviews revealed that Case 1 had recent sexual contact with a foreign national (Case 2), who had arrived in Yiwu in late July from Tanzania via Guangzhou. Case 2 reported multiple sexual encounters with local individuals between July 23 and August 7. During the subsequent investigation, we defined a "high-risk sexual network" as individuals connected through recent anonymous or casual sexual relationships occurring within the 21-day MPXV incubation period. Network membership was traced through: 1) standardized epidemiological interviews with cases and contacts, including sexual-history modules; 2) verification of travel, residence, and activity patterns via immigration and public-security databases, supplemented with mobile-signaling data when available; and 3) venue-based tracing conducted with community health service centers and managers of high-risk venues (bars, clubs, guesthouses). These findings confirmed that the outbreak originated from an imported case and spread within a confined high-risk sexual network (Figures 1 and 2).

We established a working case definition as any individual presenting with acute rash, fever, or lymphadenopathy, with documented epidemiologic linkage to a confirmed or probable case in Yiwu since July 21, and laboratory confirmation of MPXV via PCR from lesion, oropharyngeal, or other clinical samples. Active case finding encompassed reviewing outpatient and inpatient records in dermatology, infectious disease, and fever clinics; systematically interviewing identified contacts; and collaborating with community health stations. Supplementary Table S1 (available at https://weekly.chinacdc.cn/) presents a detailed chronology of case detection, reporting, laboratory confirmation, and public health

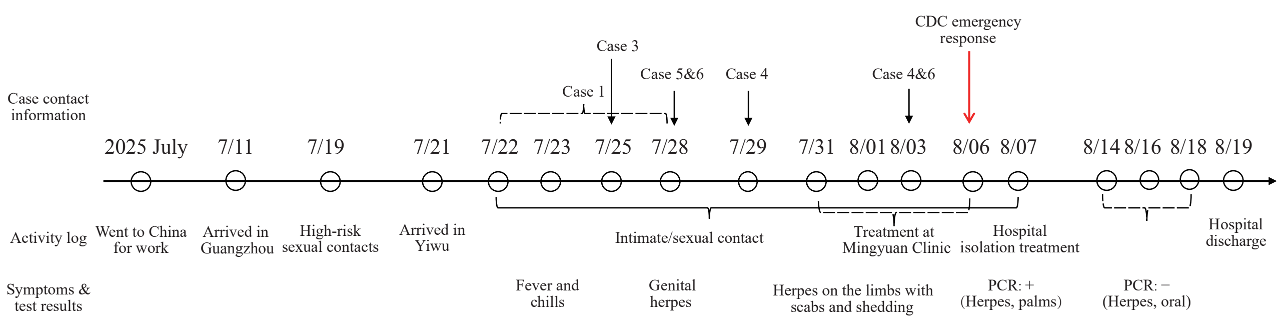


FIGURE 1. Travel and movement patterns of Case 2 confirmed by public security and immigration data — Yiwu City, July–August, 2025.
Abbreviation: PCR=polymerase chain reaction.

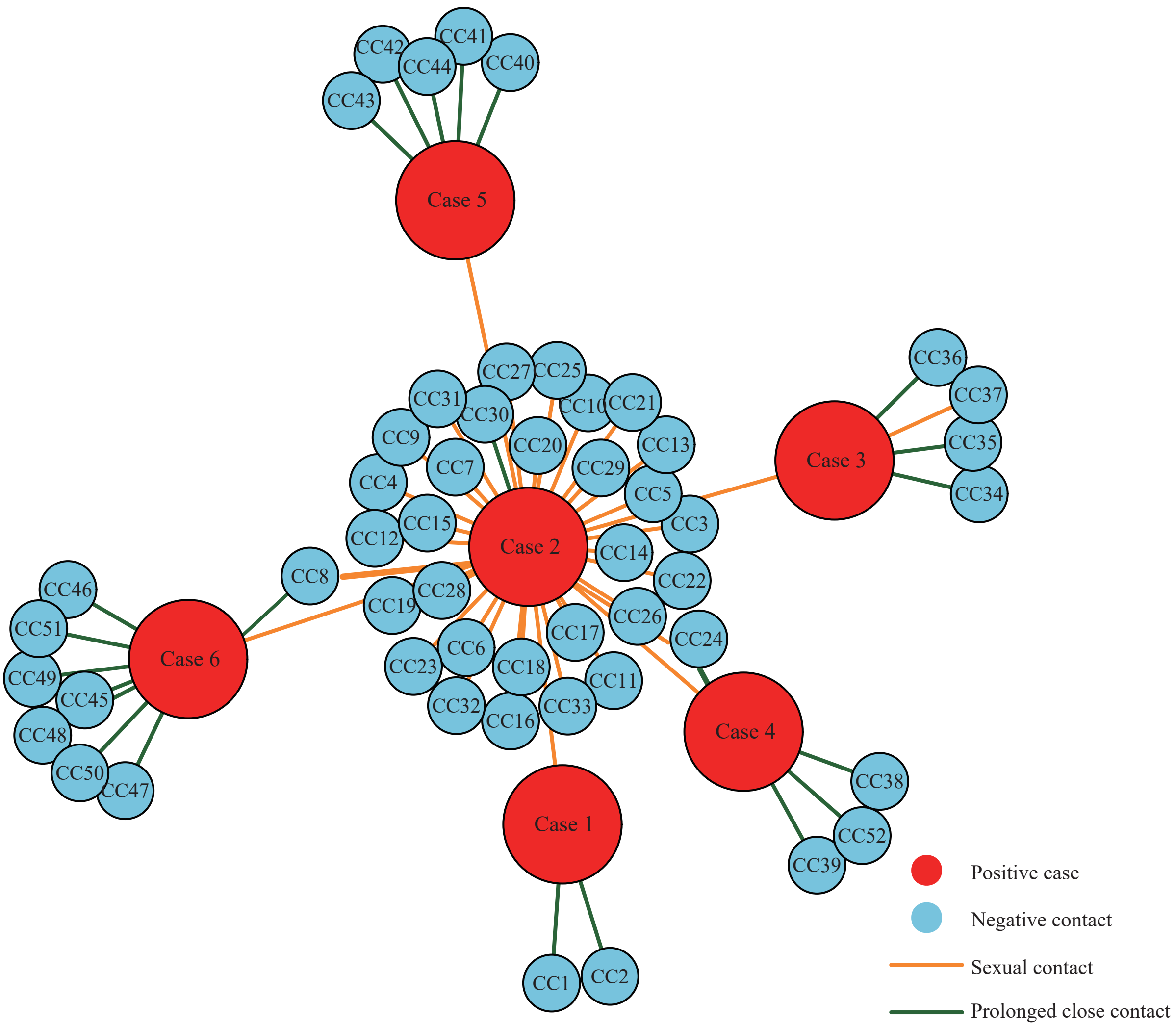


FIGURE 2. Schematic diagram of the transmission chain and contact tracing for the mpox Clade Ib cluster — Yiwu City, July–August 2025.
Note: This figure illustrates the 6 laboratory-confirmed mpox cases and all 52 core CCs identified during the investigation. The 52 core contacts comprise 31 individuals with documented sexual contact (either with the imported source Case 2, which represents the cluster's epidemiological origin, or with other confirmed cases) and 21 individuals with non-sexual prolonged close contact (including shared living spaces and repeated in-person interactions). All core close contacts tested negative for mpox throughout the 21-day health-monitoring period.
Abbreviation: CC=close contact.

intervention implementation, demonstrating the response's timeliness and effectiveness.

Between August 6 and 10, six laboratory-confirmed cases (Cases 1–6) were identified. The transmission network, illustrated in Supplementary Figure S1 (available at https://weekly.chinacdc.cn/), demonstrates that Cases 1 and 3–6 acquired MPXV infection through direct sexual contact with the imported source case (Case 2). Case 2, a female foreign national from Tanzania, was epidemiologically confirmed as the sole source of viral introduction into the network. Table 1 summarizes the demographic and clinical characteristics of all cases. Among the five cases with documented sexual exposure histories, the median interval from last sexual contact to symptom onset was 6 days (range: 5–10 days), consistent with established MPXV incubation periods. Cases were predominantly male (5/6) and aged 22–43 years (median 30 years); four were foreign nationals (from Tanzania, Azerbaijan, and Burkina Faso) and two were Chinese nationals. Clinically, all patients developed cutaneous rash lesions — predominantly affecting genital, perianal, and facial areas — accompanied by fever (100% of cases) and lymphadenopathy (67% of cases). Lesion counts ranged from fewer than 10 to more than 50. No severe complications or fatalities occurred.

TABLE 1. Demographic and epidemiological characteristics of mpox (Clade Ib) cases in the transmission chain — Yiwu City, July–August 2025.

Transmission branch	Case number & name	Gender	Age (years)	Nationality/ region	Yiwu status	Sexual contact period with index case	Onset of illness	Tested positive
Top node (index)	Case 2	Female	27	Tanzanian	Entered Guangdong (July 11 2025); Arrived in Yiwu (July 20 2025)	N/A (Index Case)	July 23	August 7
Branch 1	Case 1	Male	37	Azerbaijani	Arrived in Yiwu (July 8 2025)	July 22–July 28	August 1	August 6
Branch 2	Case 3	Male	31	Xinjiang, China	Permanent resident in Yiwu	July 25	July 31	August 7
Branch 3	Case 4	Male	22	Burkinabe	Arrived in Yiwu (May 9 2025)	July 29, August 3	August 6	August 8
Branch 4	Case 5	Male	42	Xinjiang, China	Traveled between Jiangsu and Yiwu	July 28	August 2	August 8
Branch 5	Case 6	Male	38	Tanzanian	Permanent resident in Yiwu	July 28, August 3	August 3	August 8

All patients received treatment at designated hospitals — five in Yiwu and one in Changzhou — where they received symptomatic supportive care, including antipyretics and topical treatments. Isolation was maintained until complete lesion crusting occurred and PCR testing of lesion and oropharyngeal samples returned negative results.

In total, 52 core close contacts and 38 general contacts were identified and placed under 21-day health monitoring. The 52 core close contacts included sexual partners of confirmed cases, household members, and individuals with prolonged close-range exposure; the 38 general contacts comprised primarily healthcare workers from clinics and hospitals where confirmed cases sought medical care. All contacts remaining in Yiwu underwent daily temperature and symptom monitoring for 21 days, with no mpox-compatible symptoms reported during the surveillance period.

Laboratory Testing

Laboratory confirmation for all cases relied on PCR testing of lesion and oropharyngeal specimens. Lesion samples consistently demonstrated higher viral loads than oropharyngeal swabs, as evidenced by lower Ct values (27.10±4.48 *vs.* 30.74±4.00). Notably, Case 6 exhibited positive oropharyngeal PCR results prior to rash onset, with a Ct value of 25.29, confirming that oropharyngeal viral shedding can occur before cutaneous lesion development.

Environmental investigations were conducted from August 7 to 9 in the residences of five cases and on selected personal items. Forty-three environmental surface swabs were collected from bedding, towels, bathroom fixtures, door handles, and electronic devices. The detection results of MPXV nucleic acid in different surface samples are presented inSupplementary Table S2 (available at https://weekly.chinacdc.cn/). Twenty-seven samples (62.8%) tested positive for MPXV DNA by PCR, with the highest positivity rates observed on bedding and bathroom surfaces. These findings revealed substantial and heterogeneous MPXV contamination throughout home environments, emphasizing the critical importance of comprehensive disinfection protocols.

The comprehensive investigation, supported by epidemiological and laboratory evidence, confirmed that this Yiwu cluster originated from an imported mpox case and subsequently spread within a defined high-risk network primarily through intimate and sexual contact.

PUBLIC HEALTH RESPONSE

Local health authorities and CDC teams implemented a comprehensive, multi-agency coordinated response to contain the mpox outbreak. Key interventions included: 1) Immediate isolation of confirmed cases in designated hospitals with strict infection prevention and control protocols, coupled with timely clinical management; 2) Systematic environmental disinfection of case residences and personal belongings using chlorine-based disinfectants, followed by verification testing to confirm decontamination effectiveness; 3) Enhanced active surveillance for rash and vesicular illnesses across key healthcare settings, including fever clinics, dermatology departments, and community health centers; and 4) Targeted risk communication and health education delivered in both Chinese and relevant foreign languages, with proactive outreach to high-risk social and sexual networks to promote early care-seeking behaviors and prevention practices.

These coordinated measures successfully interrupted further transmission. No additional cases were detected beyond the identified cluster, and all contacts completed the 21-day medical observation period without developing mpox-compatible symptoms, ultimately leading to formal closure of the outbreak investigation.

DISCUSSION

This investigation documented a cluster of six laboratory-confirmed mpox cases (Clade Ib) in Yiwu City, Zhejiang Province, with one imported source case and five epidemiologically linked secondary cases. This represents the first documented Clade Ib MPXV transmission cluster in Yiwu, China, driven primarily by sexual contact within high-risk networks. The importation pathway is confirmed by the travel history of the source case (Case 2), who arrived from Tanzania — a country with documented endemic circulation of Clade Ib MPXV (*6*). The outbreak occurred in an international trade hub characterized by substantial cross-border mobility, diverse sexual networks, and dense population settings, creating optimal conditions for rapid person-to-person transmission. Notably, four of the six cases were foreign nationals, highlighting the critical role of global travel in introducing mpox into non-endemic regions.

The epidemiologic pattern reflects recent global outbreaks where close skin-to-skin and sexual contact serve as the primary transmission routes (*7*). However, the involvement of heterosexual commercial sex in this cluster expands the recognized transmission contexts in China. The attack rate within this defined sexual network was high, consistent with reports documenting efficient Clade Ib transmission within sexual networks from other regions (*7–8*). Unlike scenarios involving household transmission, particularly to children in endemic areas (*7–8*), we observed no secondary transmission to household contacts despite extensive environmental contamination. This suggests that while highly transmissible through intimate contact, the effective reproduction number (R-eff) in non-intimate household settings may be lower, potentially influenced by viral load, contact type, and timely decontamination efforts. This contrast emphasizes the importance of context-specific transmission risk assessment. Notably, one case (Case 6) yielded a positive oropharyngeal real-time PCR result before rash onset, suggesting potential pre-symptomatic viral shedding. While this single observation is intriguing, it cannot confirm pre-symptomatic transmission. However, it aligns with clinical guidelines indicating that some cases may be infectious 1–4 days before symptom onset (*9*). This finding warrants further investigation and underscores the need to re-examine current contact tracing and isolation protocols to account for potential pre-symptomatic transmission (*10*).

Extensive environmental contamination — identified through PCR positivity on bedding, household surfaces, and personal hygiene items — demonstrates widespread viral DNA distribution and highlights the potential for indirect transmission via contaminated objects (*8*). This finding reinforces the critical importance of comprehensive environmental decontamination in outbreak response protocols (*11*). However, it is essential to recognize that PCR detection of viral DNA does not necessarily indicate the presence of viable, infectious virus, which requires confirmation through cell culture methods. The rapid containment achieved through coordinated interventions — including timely case isolation, systematic contact tracing, targeted health education, and thorough environmental disinfection — demonstrates the effectiveness of multi-agency public health responses.

This investigation has several important limitations. First, the small outbreak size limits our statistical power to draw definitive conclusions, particularly regarding the significance of pre-symptomatic oropharyngeal viral shedding observed in one case. Second, the inherently clandestine and anonymous nature of high-risk sexual networks likely resulted in incomplete case detection and underreporting, suggesting our investigation may not have captured the full transmission network. Third, without viral culture performed on environmental samples, we cannot confirm whether the detected viral DNA represented infectious virus, potentially leading to overestimation of fomite transmission risk. Future investigations with larger sample sizes and systematic viral culture integration are needed to validate these preliminary findings.

Despite these limitations, this outbreak underscores the necessity of proactive surveillance within high-risk sexual and social networks, culturally appropriate risk communication strategies, and the integration of environmental and genomic data into comprehensive outbreak management. Additionally, the observed pre-symptomatic oropharyngeal viral shedding warrants further investigation to refine diagnostic protocols and optimize prevention strategies. Sustained vigilance, rapid detection capabilities, and cross-sectoral collaboration remain essential for preventing future introductions and limiting the spread of MPXV (Clade Ib) in China's high-mobility urban centers.

Ethical statement: Approval from the Ethics Committee of Jinhua Center for Disease Control and Prevention, China (approval number: 2025-22).

Conflicts of interest: No conflicts of interest.

doi: 10.46234/ccdcw2025.221

[#] Corresponding authors: Jiming Sun, jmsun@cdc.zj.cn; Jun Jiang, junjiang3176@163.com.

[1] Jinhua Center for Disease Control and Prevention, Jinhua City, Zhejiang Province, China; [2] Yiwu Center for Disease Control and Prevention, Yiwu City, Zhejiang Province, China; [3] Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou City, Zhejiang Province, China; [4] Center for Global Public Health, Chinese Center for Disease Control and Prevention, Beijing, China.
[&] Joint first authors.

Submitted: August 26, 2025
Accepted: September 29, 2025
Issued: October 10, 2025

REFERENCES

1. Jezek Z, Grab B, Szczeniowski MV, Paluku KM, Mutombo M. Human monkeypox: secondary attack rates. Bull World Health Organ 1988;66 (4):465 – 70.
2. Adepoju P. Mpox declared a public health emergency. Lancet 2024;404 (10454):e1 – 2. https://doi.org/10.1016/S0140-6736(24)01751-3.
3. Sun JM, Zhou L, Wu BB, Chen ZW, He YD, Wang XX, et al. Characteristics of the first confirmed case of human infection with Mpox virus clade Ib in China. Nat Commun 2025;16(1):4888. https://doi.org/10.1038/s41467-025-60217-2.
4. Allan-Blitz LT, Gandhi M, Adamson P, Park I, Bolan G, Klausner JD. A position statement on Mpox as a sexually transmitted disease. Clin Infect Dis 2023;76(8):1508 – 12. https://doi.org/10.1093/cid/ciac960.
5. Titanji BK, Hazra A, Zucker J. Mpox clinical presentation, diagnostic approaches, and treatment strategies: a review. JAMA 2024;332(19): 1652 – 62. https://doi.org/10.1001/jama.2024.21091.
6. UK Health Security Agency. Mpox: affected countries. UK Health Security Agency. 2025. https://www.gov.uk/guidance/clade-i-mpox-affected-countries. [2025-08-26].
7. Ai L, Cui XY, Zhou YQ. Clade Ib Mpox virus: how can we respond? China CDC Wkly 2025;7(29):979-84. http://dx.doi.org/10.46234/ccdcw2025.165.
8. Srivastava S, Laxmi, Sharma K, Sridhar SB, Talath S, Shareef J, et al. Clade Ib: a new emerging threat in the Mpox outbreak. Front Pharmacol 2024;15:1504154. https://doi.org/10.3389/fphar.2024.1504154.
9. National Administration of Disease Prevention and Control. Plan for the prevention and control of Monkeypox. National Administration of Disease Prevention and Control. 2023. https://www.chinacdc.cn/jkyj/crb2/yl/hd/jswj_hd/202409/W020240906514307430147.pdf. [2025-08-26].
10. Rahman MM. Systematic Review of the Epidemiology, disease severity and preventive measures of Mpox disease: a public health guideline for densely-populated countries. F1000Research 2024;13:590. http://dx.doi.org/10.12688/f1000research.150476.1.
11. Atkinson B, Gould S, Nicholls I, Manzoor KN, Smith J, Hindle AJ, et al. Air and surface sampling for clade Ib monkeypox virus in United Kingdom hospitals, 2024 to 2025. Euro Surveill 2025;30(19):2500288. https://doi.org/10.2807/1560-7917.ES.2025.30.19.2500288.

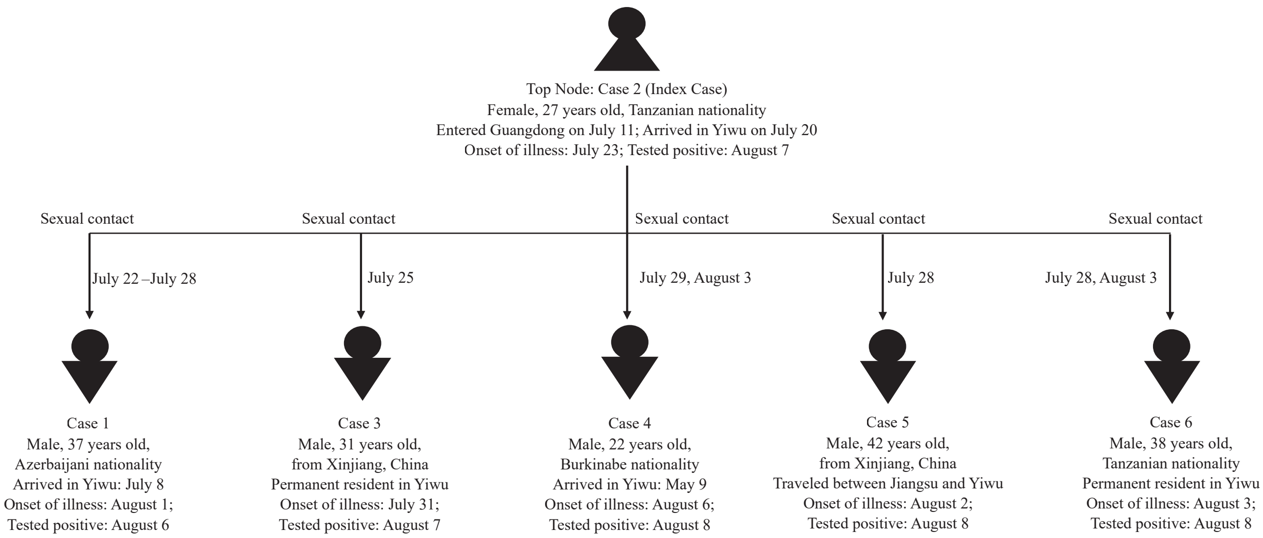
SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. Chronological timeline of events and public health interventions during mpox Clade Ib outbreak —Yiwu City, July–August 2025.

Date	Event/node	Public health intervention
July 11	Imported source case (Case 2) entered China via Guangzhou	Not recognized (incubation period)
July 21	Case 2 arrived in Yiwu	Not recognized (incubation period)
July 23–30	Case 2 developed symptoms	Self-medication, no medical visit
July 31–Aug 4	Case 3 developed symptoms, multiple hospital visits	Not recognized, symptomatic treatment
July 31–Aug 5	Index case (Case 1) developed symptoms, multiple clinic/hospital visits	Not recognized, symptomatic treatment
July 31–Aug 6	Case 2 visited clinics repeatedly	Not recognized, symptomatic treatment
Aug 2–5	Case 5 developed symptoms	No medical visit
Aug 3–8	Case 6 developed fever, cough, lymphadenopathy	No medical visit
Aug 6–7	Case 4 developed symptoms	Self-medication, no medical visit
Aug 6	Case 1 detected and reported by Yiwu Central Hospital; Yiwu CDC PCR positive; patient isolated the same night	Emergency response activated; case transfer, lab testing, preliminary epidemiologic investigation
Aug 7	Jinhua CDC confirmed Case 1; Case 2 identified and confirmed via investigation; Case 3 revisited and tested positive; 21 field teams established	Provincial–municipal expert guidance; “three-sector” joint tracing; contact tracing and management; case isolation; environmental sampling and disinfection
Aug 8	Cases 4, 5, and 6 identified as close contacts, PCR positive	Contact tracing and management; case isolation; environmental sampling and disinfection
Aug 9	Last case (Case 6) transferred for management	Completion of centralized isolation and treatment
Aug 10–12	21-day monitoring of close and general contacts initiated	Continued tracing, testing, and health monitoring; environmental sampling and disinfection
Aug 19–22	Discharge of hospitalized cases in Yiwu	Medical treatment with psychosocial support
Aug 30	Completion of 21-day monitoring of close contacts	No new cases; outbreak under control
Sep	Outbreak report finalized	Evaluation of interventions; recommendations for improvement

SUPPLEMENTARY TABLE S2. Environmental detection results for Mpox virus DNA from case residences — Yiwu city, July–August 2025.

Object	Sample type	Total number of samples collected	Number of positive samples	Names of positive sites	Names of negative sites
Case 1	External environment of residence	2	1	Staircase	Door handle
	Internal environment of residence	8	8	Pillow, quilt, stool, sofa, towel, desktop, toilet, washbasin	/
	Personal belongings	/	/	/	/
Case 2	External environment of residence	3	1	Door handle	Staircase, door handle on the first floor
	Internal environment of residence	7	5	Pillow, quilt, TV cabinet, sofa, desktop	Toilet faucet, towel
	Personal belongings	1	1	Mobile phone	/
Clinics visited by Case 1, Case 2, and coffee shop they frequented	/	10	0	/	Stool in coffee shop, table in coffee shop, Weiqiang Clinic (stool, infusion stool, infusion sofa, table), Jiaming Clinic (stool, table), Zhong Xiaorong Clinic (door handle, desktop in clinic)
Case 3	External environment of residence	2	1	Door handle (inside and outside)	Door handle of company (outside)
	Internal environment of residence	5	3	Door handle (master bedroom), bedside switch and charger, desktop and kettle handle	Door handle of company (inside)
	Personal belongings	0	0	/	/
Case 4	External environment of residence	1	1	Door handle	/
	Internal environment of residence	5	5	Refrigerator, wardrobe, door handle, etc.	/
	Personal belongings	0	0	/	/
Case 6	External environment of residence	1	0	/	Door handle
	Internal environment of residence	7	0	/	Towel, bed, etc.
	Personal belongings	1	1	Mobile phone	/
Total	Smear samples from living environment and personal belongings	43	27	/	/
	External environment (nonresidence)	10	0	/	/



SUPPLEMENTARY FIGURE S1. Transmission chain of the Mpox (Clade Ib) outbreak —Yiwu City, July–August 2025.

Indexed by Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), PubMed Central (PMC), Scopus, Chinese Scientific and Technical Papers and Citations, and Chinese Science Citation Database (CSCD)

The inauguration of *China CDC Weekly* is in part supported by Project for Enhancing International Impact of China STM Journals Category D (PIIJ2-D-04-(2018)) of China Association for Science and Technology (CAST).

CCDC weekly

Vol. 7 No. 41 Oct. 10, 2025

Responsible Authority

National Disease Control and Prevention Administration

Sponsor

Chinese Center for Disease Control and Prevention

Editing and Publishing

China CDC Weekly Editorial Office
No.155 Changbai Road, Changping District, Beijing, China
Tel: 86-10-63150501, 63150701
Email: weekly@chinacdc.cn

CSSN

ISSN 2096-7071 (Print)
ISSN 2096-3101 (Online)
CN 10-1629/R1