ISSN 2096-7071 (Print)
ISSN 2097-3101 (Online)
CN 10-1629/R1

CHINA CDC WEEKLY

Vol. 8 No. 20 May 15, 2026

中国疾病预防控制中心周报(英文)

MONKEYPOX ISSUE

ISSN 2096-7071

China CDC Weekly

Editorial Board

Editorial Office

Cover Image: credited from the designer Jiaying Wang.
This week's issue was organized by Guest Editor Zhongjie Li.

Review

A Review of Epidemiological Modeling Studies on Monkeypox

Jiahui Li[1,&]; Qixuan Luo[2,&]; Emiliano Beltran[3,4]; Qiuping Chen[1,2,5]; Sandra Perez[6,7]; Roger Frutos[1,2,5,8,9]; Yanhua Su[1,2]; Ziyan Liu[10]; Jia Rui[11,12]; Tianmu Chen[1,2]; Kaiwei Luo[10,#]; Zeyu Zhao[1,2,4,13,#]

ABSTRACT

Between 2022 and 2025, nearly 130,000 confirmed monkeypox (mpox) cases, including over 280 deaths, were reported to the World Health Organization (WHO) from 130 countries and territories, prompting the WHO to declare it a public health emergency of international concern on two occasions. Hence, to enable a scientific approach to prevention and control, transmission dynamics through mathematical modeling must be elucidated. Through a comprehensive literature review on the modeling of mpox transmission dynamics, this study indicates that existing research primarily extends the susceptible-infectious-recovered (SIR) modeling framework and largely focuses on the 2022 global outbreak. The analysis revealed significant variations in mpox virulence, particularly dependent on the subtypes, and variations according to the descending order of clades: Ib>Ia>IIa>IIb. Interpersonal transmission capacity was notably higher for clade II than for clade I, highlighting geographical disparities, with the highest transmission capacity in the Americas, moderate in Europe and Oceania, and the lowest in Asia. In contrast, Africa maintained consistently low but non-declining transmission levels. Furthermore, the study confirmed significant co-infection patterns between mpox and sexually transmitted diseases, such as Human Immunodeficiency Virus (HIV) and syphilis, along with evidence of synergistic transmission interactions among multiple pathogens. This study provides crucial modeling evidence for mpox regulation. By quantifying high-risk populations, evaluating intervention effectiveness, and identifying the risks associated with subtypes and geographic areas, this study offers a comprehensive reference for understanding the interface between model complexity and practical application. These advancements have enhanced the strategic focus and precision of public health responses during outbreaks.

Monkeypox (mpox) is a zoonotic infectious disease caused by the monkeypox virus (MPXV), with clinical manifestations similar to those of smallpox but with lower incidence and mortality (*1*). MPXV is classified into clades I (formerly the Central African or Congo Basin clade, subdivided into Ia and Ib) and II (formerly the West African clade, subdivided into IIa and IIb). Transmission occurs via direct contact with infectious lesions, body fluids, respiratory droplets, or contaminated fomites (*2*). From 1970 to 2017, mpox was endemic primarily in Central and West Africa, predominantly via animal-to-human transmission. Sporadic cases in non-endemic regions were largely related to travel or animal importation (*3*–*4*). However, human-to-human transmission has increased following the 2017 Nigerian outbreak (*5*). In May 2022, a large global outbreak occurred, spreading primarily through sexual contact networks (*6*), with cases reported among men who have sex with men (MSM).

The epidemiological modeling framework for mpox has evolved over the decades, with research on transmission dynamics among humans and between humans and animals, forming a relatively complete theoretical system. This study was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews guidelines (*7*). The literature search covered the China National Knowledge Infrastructure (CNKI), Web of Science, and PubMed for the period spanning 3 years following the global spread of mpox in 2022 (2022–2025). The focus was on the literature employing mathematical modeling methods to study the transmission dynamics models of orthopoxvirus (mpox) among humans or between humans and animals. The searches were conducted on December 25, 2025, using keywords including "mpox," "Ia/Ib/IIa/IIb," "dynamic" and "model." "Monkeypox" was also included as a keyword for the 2022 publications to account for the WHO nomenclature change in November 2022. This study did not include

gray literature, and all analyses were based on published peer-reviewed academic literature. No geographic or language restrictions were applied.

The systematic search yielded 11,824 relevant records. After deduplication, 5,479 duplicate records were removed. The remaining records were independently screened by two reviewers in a two-round process and discrepancies were resolved by consensus. The first round, based on title and abstract review assessing the application of mathematical compartmental models, identified 122 articles for full-text assessment. During the second round of full-text screening, 39 articles that did not analyze the epidemic transmission potential and 4 dissertations were excluded. Ultimately, 79 studies were included in the systematic review of compartmental structures, parameter settings, and dynamic characteristics of the models (Figure 1).

COMPARTMENTAL MODEL STRUCTURE

Since Kermack and McKendrick introduced the classic susceptible-infectious-recovered (SIR) compartmental model in 1927 (*8*) and elucidated the threshold theory of infectious disease dynamics in 1932 (*9*), mathematical compartmental models have become essential tools for evaluating epidemic prevention and control strategies. The study of mpox transmission dynamics still employed the basic SIR model to depict the fundamental transmission process of the disease (*10–13*). Based on the natural history of mpox virus infection, most studies have introduced an exposed compartment E (exposed) to construct the susceptible-exposed-infectious-recovered (SEIR) model (*14–18*), whereas a few have incorporated an asymptomatic infectious compartment A into the

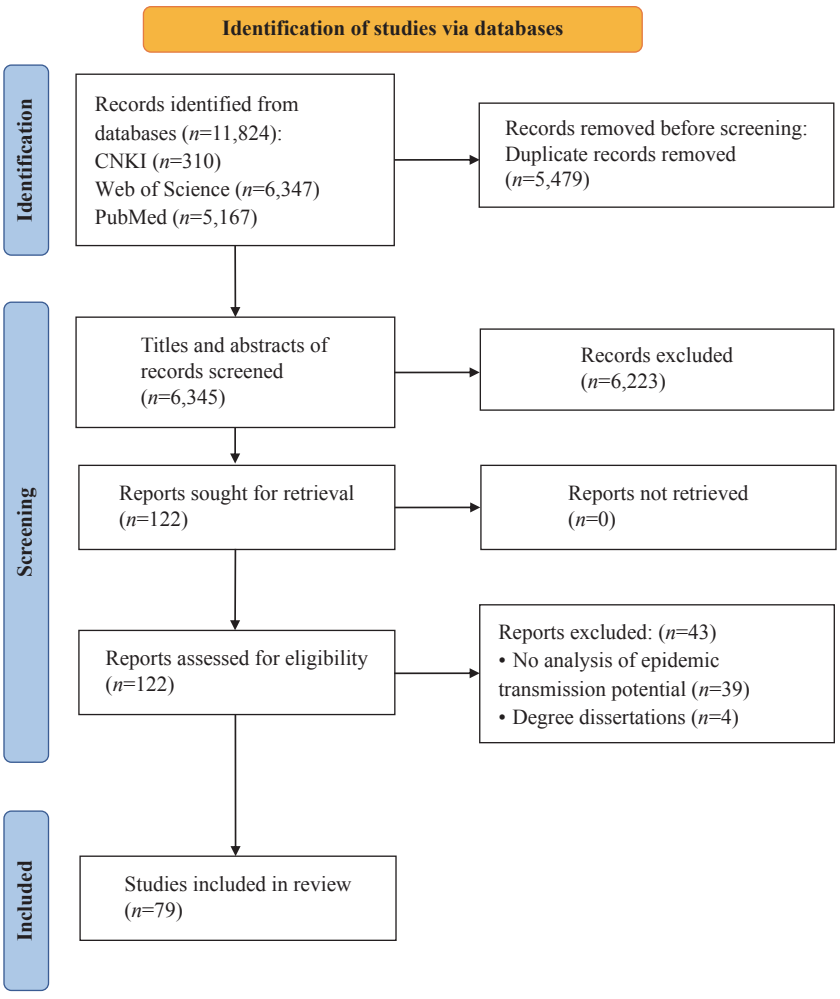


FIGURE 1. Flowchart of literature screening and selection results.
Abbreviation: CNKI=China National Knowledge Infrastructure.

model framework (*19–20*) as a possible state following the exposure period, thereby providing a more comprehensive representation of the clinical spectrum and transmission characteristics of mpox. Given the zoonotic nature of mpox, numerous models have also incorporated animal populations (e.g., rodent or murine species), typically with simplified compartmental structures, such as susceptible-infectious (SI) (*21–29*), susceptible-exposed-infectious (SEI) (*30–38*), SIR (*39–40*), and SEIR (*41–48*). To distinguish them from human compartments, animal compartments are often labeled with subscripts, such as m, r, or a. Infected animal compartments (I_m, I_r, and I_a) serve as sources of infection, participating in both transmission within susceptible animal populations (S_m, S_r, and S_a) and cross-species transmission to susceptible humans (S_h). A few models consider reverse transmission from infected humans (I_h) to susceptible animals (*25,34,44,46,49*). Through a systematic review of the relevant models, this study summarizes a schematic diagram of the mpox transmission patterns (Figure 2), revealing that the evolution of model frameworks has always been closely aligned with their transmission characteristics.

To better capture the complexities of transmission, several studies have extended model frameworks by incorporating population heterogeneity and dynamic control measures. Spatial and demographic heterogeneity are addressed by expanding compartments with subscript i, representing either different countries to model the effect of cross-border mobility on epidemic spread (*50–51*) or different age groups to analyze variations in disease prevalence across age structure (*13,52–53*). Furthermore, given the high incidence of mpox among MSM, some models further stratify them based on sex or sexual behavior risk, dividing it into male (subscript m) and female (subscript f) groups (*54*) or into high-risk (subscript h) and low-risk (subscript l) groups (*50,55–60*), to construct coupled multi-group models. Vaccination, the most frequently considered intervention, is typically modeled before exposure. This is achieved by establishing dedicated compartments for vaccinated (V) or protected (P) (*60–64*) or by further subdividing the susceptible population into vaccinated (S_v) and unvaccinated (S_u) subgroups (*65*), representing the transition from susceptible to an immune state. Regarding the sequence of vaccination and isolation, some studies have also incorporated a post-exposure vaccination compartment (Q_v) (*50*). Models vary in their assumptions regarding vaccine-induced immunity, ranging from complete protection

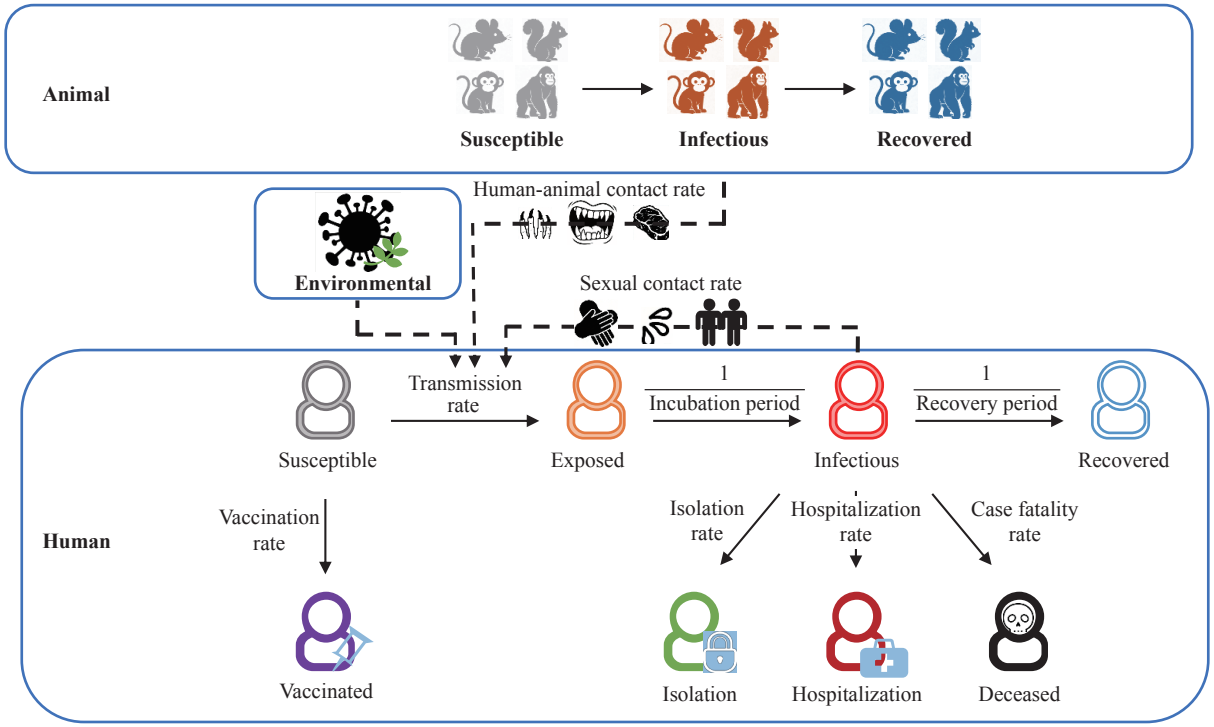


FIGURE 2. Schematic diagram of the conceptual framework and key components for mpox transmission dynamics modeling.
Abbreviation: mpox=monkeypox.

(*57,63–64,66*) to accounting for breakthrough infections (*19,50,52,58–60,67–68*) and waning immunity (*18,21*). Furthermore, isolation (Q) (*19,38,50,56,60–62,69–71*) or hospitalization (H) (*60,68*) compartments are included to manage symptomatic cases; a deceased compartment (D) (*41,62–64,67,72*) captures mortality outcomes; and an environmental contamination compartment (C) (*73–74*) simulates pathogen persistence, dissemination in the environment, and its role as a potential source of infection.

Owing to the temporal overlap of the mpox outbreak with the coronavirus disease 2019 (COVID-19) pandemic in 2022, some current modeling studies include analyses of the transmission dynamics of mpox co-infection with COVID-19 (*75*). Moreover, given that the epidemiological characteristics of mpox is primarily transmitted through sexual contact, co-infection with sexually transmitted pathogens such as HIV (*54,76–79*) and *Treponema pallidum* (*80*) has also become a significant focus of modeling research. Among these, models of mpox and HIV co-infection are relatively more common. These models typically construct three parallel transmission chains: monkeypox virus-only infection, HIV-only infection, and co-infection with both pathogens, with the susceptible population generally treated as a unified group. Common compartments include individuals infected only with mpox (I_m), individuals infected only with HIV (I_h), individuals infected with HIV receiving antiretroviral therapy (ART) (I_h^T), and individuals co-infected with both viruses (I_{mh} and co-infected individuals receiving ART, I_{mh}^T). As the classification of infections expands, the compartmental structure and dimensionality of the models increase significantly. Although this better reflects real-world complexity, it also significantly increases the difficulty of analyzing model dynamics.

MODEL PARAMETERIZATION

As core parameters of compartmental models, the values of the transmission rate coefficient (β), latency period, and recovery period vary depending on the study context, population structure, and transmission routes. Global data indicate an average human-to-human transmission rate of approximately 0.09 per day (*64*). The rate is significantly higher in high-risk populations, predominantly among MSM (*50,55,58,67–68,75*), with a per-sexual contact transmission rate of 0.24/day in Canada (*56*) and 0.34/day in the United States of America (USA) (*81*) in 2022. In contrast, the non-sexual contact rate in low-risk populations is only about one-thousandth of the sexual contact rate (*82*), and the estimated β value for the environmental transmission route is as low as 7.70×10^{-8} (*73*). The average incubation period is primarily considered as 8.5 (*11,51,56,58*) and 13 (*50,64*) days. Africa Centers for Disease Control and Prevention supports an average of 13 days (*64*), whereas Western countries adopt a mean of 8.5 days (*51,56,58*), possibly reflecting regional differences in population immunity. The prodromal period lasts 1–4 days (*11,83–84*); though asymptomatic, the mean infectiousness coefficient is between 0.25 and 0.51 (*82*).

For recovery, without isolation, untreated infected individuals have a mean recovery period of 21 days (*20,50,56,60,63–64,74*), with no significant differences in recovery rates across populations (*81*). The weekly recovery rate is 1.4×10^{-3} for both children/adolescents and adults (*13*) and similar for symptomatic and asymptomatic cases (*20,56,64*). Vaccination shortens recovery (*20,25,61,67,74*), and pre-exposure coverage of 20%–50% in high-risk populations yields major control benefits (*28,61*). Two-dose recipients recover approximately 30% faster (*85*), whereas vaccine waning rates range from 3.17×10^{-3} to 2.04×10^{-2} per day (*19*). Protection lasts approximately 5 years for a single dose and approximately 10 years for two doses, similar to natural immunity (*67*). However, modeling evidence confirms that residual immunity from historical smallpox vaccination alone is insufficient to prevent or contain mpox outbreaks (*39*). In fact, vaccination and isolation exhibit synergistic effects. Specifically, at ≤20% vaccination coverage, a 40% isolation rate is required to control transmission, whereas at 30% coverage, a 20% isolation rate remains effective (*61*).

For zoonotic diseases, the human-to-animal contact rate is a key parameter in models of initial outbreaks, areas with a clear animal exposure risk, or transmission chains traceable to animal contact (*32-33,44–45,48*). Its value varies according to regional habits, ecology, and host distribution. In endemic regions of Africa, where rural residents hunt, consume, or keep wild animals, the estimated human-to-animal contact rates are significantly higher than in non-endemic areas (*34,44–45,48*), ranging from 2.50×10^{-4} to 2.70×10^{-2} (*33,40*). The per-contact transmission probability from rodents to humans is 0.13 in the Democratic Republic

of the Congo (DRC) and as high as 0.85 in Nigeria (*45*). In contrast, outbreaks in Europe and North America are primarily driven by human-to-human transmission (*26*). During the 2022 outbreaks, the transmission rate associated with human-to-animal contact was 5.00×10^{-4} in Spain and 3.30×10^{-4} in Italy. Sensitivity analysis identified the animal-to-animal transmission rate and human-to-animal contact rate as critical parameters influencing R_0. Interventions targeting animal reservoirs and reducing human-to-animal contact rates could potentially reduce the number of infections by over 80% (*44*).

HIV and *T. pallidum*, as significant co-infecting pathogens in mpox transmission, influence mpox transmission dynamics primarily via susceptibility (*77–78*), infectiousness (*77–78*), and response to interventions (*77*), especially in high-risk populations such as MSM (*77,86*). Regarding susceptibility parameters, HIV-infected individuals have a susceptibility adjustment factor for mpox of 1.5, which has remained at 1.3, even for those receiving ART (*77–78*). Regarding the infectiousness parameters, the mpox infectiousness adjustment factor for HIV-infected individuals receiving ART is typically <1 (*54,77*), indicating that ART moderately reduces the transmission potential of co-infected individuals. Syphilis infection increases susceptibility to mpox to a lesser extent than HIV infection, with an adjustment factor typically set at 1.2. Regarding transmission dynamics, the basic transmission rate is approximately 6.50×10^{-2} for syphilis and approximately 2.01×10^{-1} for mpox. Co-infection with both pathogens exerts a synergistic effect, significantly enhancing the transmission potential compared with mono-infection (*87*). The selected parameter values reported in the literature are summarized in Table 1.

TRANSMISSION POTENTIAL

Transmissibility and virulence of mpox viruses vary by clade. Clade I, which is endemic to the Congo Basin, is characterized by high fatality rates (up to 10%) and zoonotic transmission with limited human-to-human spread (*88*). Subtypes Ia and Ib harbor a high proportion of apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3)-driven cells (68% and 72%, respectively), suggesting increased adaptability and transmission potential (*89*). Notably, since September 2023, the newly identified Ib subtype among sex workers in the DRC exhibited high fatality rates (approximately 5.3% overall and 8%–10% in children) (*90–93*), providing preliminary evidence of increased virulence and transmission efficiency. In contrast, clade II is predominantly

TABLE 1. Selected parameters for mpox compartmental models from the literature.

Parameter	Units	Values	Reference
Human epidemiological parameters for mpox			
Human-to-human transmission rate	$Days^{-1}$	0.09 (international average value)	(*64*)
		0.24 (sexual, Canada)	(*56*)
		0.34 (sexual, USA)	(*81*)
Transmission rate coefficient	N/A	7.70×10^{-8} (environmental transmission)	(*73*)
		0.25–0.51	(*82*)
Incubation period	Days	8.5	(*11,51,56,58*)
		13 (average)	(*50,64*)
Prodromal period	Days	1–4	(*11,83–84*)
Recovery period	Days	21	(*20,50,56–60,63–64,74*)
Recovery rate	$Weeks^{-1}$	1.4×10^{-3}	(*13*)
Vaccine waning rate	$Days^{-1}$	3.17×10^{-3}–2.04×10^{-2}	(*19*)
Animal transmission mpox infection parameters			
Human-to-animal contact rate	N/A	2.50×10^{-4} –2.70×10^{-2}	(*33,40*)
Per-contact transmission probability	N/A	0.13 (DRC)	(*45*)
		0.85 (Nigeria)	(*45*)
Human-animal contact rate	N/A	5.00×10^{-4} (Spain)	(*26*)
		3.30×10^{-4} (Italy)	(*26*)

prevalent in West Africa, demonstrating higher human-to-human transmissibility and lower fatality rates of approximately 1% *(88)*. Clade IIa, which shares approximately 95% nucleotide sequence homology with clade I, follows a zoonotic transmission pattern (*88*). Clade IIb, the primary subtype responsible for a global outbreak in 2022, combines high human transmissibility with relatively low virulence (*88*,*93*).

On May 7, 2022, the United Kingdom (UK) reported its first imported case, after which the outbreak rapidly spread to multiple countries in Europe and the Americas. During the early phase, transmission was most rapid in Europe, Oceania, North America, and Latin America, with the median basic reproduction number (R_0) reaching 7.02 (range 7.01–7.84). Asia and Africa recorded lower values of 3.68 and 1.61, respectively. After June 23, 2022, the transmission rates declined significantly in most regions except Africa. The median R_0 value in the medium- to high-risk countries dropped to 1.35. Geographically, R_0 remained at relatively high levels in Latin America and North America (4.13 and 3.51, respectively), followed by Oceania and Europe (1.67 and 1.35, respectively), whereas Asia recorded the lowest (0.70) (*51*). By December 2022, the cumulative number of cases in non-endemic countries exceeded 80,000. Transmission during this outbreak was primarily concentrated in the MSM population. A model based on this group was fitted with a global effective reproduction number (R_{eff}) of 3.11 (*57*).

Transmission in North America occurred significantly earlier in the outbreak. In Canada, R_0 was relatively high, at R_0=1.68 before May 6, 2022 (*57*), with an estimated annual value of 2.37 (*15*,*70*). R_0 for the high-risk, low-risk, and total populations were 1.464, 0.007, and 1.461, respectively. Intervention simulations targeting the high-risk group demonstrated that condom use and abstinence would require adherence rates of 25% and 10%, respectively, to block transmission (*56*,*70*), reducing R_{eff} to 0.69 after behavioral changes (*58*). In the USA, the concurrent R_0 was 3.28 (*15*,*81*), which was significantly higher than that of Canada, with R_0 of 3.88 and 0.39 for the high-risk and low-risk groups, respectively. A two-dose vaccination campaign targeting a high-risk population (71.8% of cases) could prevent approximately 21.2% of infections, whereas behavioral changes alone could prevent 15.4% of cases. This combination exhibited synergistic effects (*20*). Transmission intensity in Latin America was high. Colombia had the highest R_0 at 3.60, with approximately 25% of the actual cases reported on average (*70*). R_0 for Brazil, Peru, and Mexico are 2.61, 2.49, and 2.51, respectively (*15*). In Peru, reported cases were 96.3% male, with a mean age of 32.6 years. MSM accounted for 75.0% of the cases, and the HIV coinfection rate reached 55.6% (*94*).

Transmission in Europe fluctuated. During the initial 2022 outbreak, Spain's R_0 was 2.05 (*57*), increasing to 3.05 for the annual estimate (*15*). Germany's R_0 before May 6, 2022, was 3.21 (*15*), with the annual estimate rising to 3.57 (*57*). The annual R_0 estimates for France and the UK were 2.77 and 2.43, respectively (*15*), whereas Portugal's estimates were lower at 1.05 (*57*). In Africa, Uganda exhibited the highest R_0 at 2.51, followed by Rwanda and Burundi with 1.35 and 1.88 (*65*). In China, an analysis of 1,433 WHO-reported cases from April 3 to September 16, 2023, yielded an initial R_0 of 1.43 (*69*). Following the isolation interventions, R_{eff} decreased to 0.34 (*69*), indicating a transmission intensity lower than the global average.

DISCUSSION

Historically, mpox outbreaks have been largely confined to Africa. However, since the rapid global spread of subclade IIb began in May 2022, the construction of scientifically sound dynamic transmission models has become essential for understanding epidemic patterns and guiding prevention and control strategies. This study systematically reviews 79 publications, tracing the evolution of models from basic SIR/SEIR frameworks to multi-layered, heterogeneous structures incorporating sexual contact networks, cross-species transmission, and multi-pathogen co-infections. It also analyzes key parameters, as well as regional and population differences in transmission potential. The findings reveal a clear trade-off between virulence and transmissibility among viral subclades; while subclade IIb exhibited attenuated virulence, its human-to-human transmissibility was markedly enhanced. Geographically, apart from traditional endemic areas in Africa, the transmission intensity follows a gradient pattern of "Americas > Europe and Oceania > Asia." This distribution is closely associated with regional variations in sexual network density among MSM, the speed of vaccination coverage, and the intensity of public health responses. This work provides a methodological foundation for rapidly modeling and early warning of future mpox outbreaks and offers a

theoretical reference for the early response to potential "disease X" scenarios with similar transmission mechanisms.

During the literature search, different modeling approaches addressed the "what" or "why" questions. However, when the research question advanced to "how to control the outbreak" and "which interventions are most effective," the unique advantages of transmission dynamic models became evident. For instance, a modeling study of the 2022 New York City outbreak subdivided the population into five age groups and seven risk categories and concluded that targeted immunization of high-risk individuals aged 35–44 significantly enhances control effectiveness (*52*). This exemplifies the key distinction between dynamic transmission models and other modeling approaches. Their core value lies in their ability to quantitatively simulate the effects of different control measures, integrate the synergistic effects of multiple interventions such as vaccination, isolation, and behavioral change, and support the dynamic adjustment of strategy parameters as the epidemic evolves.

Although dynamic models hold irreplaceable value in strategy evaluation, existing studies have significant limitations. Some have focused on population stratification within a single region, which limits their ability to capture the complexities of the global spread of mpox, escalating geopolitical conflicts, and reduced international aid (*95*). The 2025 early-phase outbreak of the Ib subclade in China following imported cases (*96*) further underscores the urgent need to shift research from localized static analysis to cross-regional dynamic modeling. A Singapore modeling study demonstrated that dynamically adjusting border strategies based on prevalence rates in countries of origin could effectively reduce importation risk by approximately 21.8% (*14*), confirming the necessity and feasibility of cross-regional modeling. However, parameter estimation often relies on assumptions or literature values without clear justification, and insufficient surveillance, together with a lack of data standardization, may lead to systematic deviations between model predictions and reality (*6*).

In future research, the design of dynamic models must strike a balance between complexity and operability, thereby enhancing the accuracy and interpretability of parameters in real-world contexts through multi-model integration. It is essential to integrate multi-regional population mobility, intervention measures, and surveillance data to construct a dynamic framework that supports differentiated cross-border prevention and control and optimizes resource allocation, thereby strengthening model adaptability and decision-support capabilities in complex, real-world outbreak scenarios. Meanwhile, interdisciplinary research should explore collaborative modeling frameworks that protect data privacy, such as federated learning, and promote their validation and application in authentic clinical environments. For China, efforts should focus on conducting timely and prospective analyses using new data and emerging outbreaks, optimizing parameter settings based on local surveillance characteristics, and developing targeted analytical frameworks suitable for both outbreak emergency response and routine surveillance scenarios, thereby better supporting localized scientific prevention and control of the epidemic.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the National Key Research and Development Program of China (2024YFC2311404), the Self-supporting Program of Guangzhou Laboratory (GZNL2024A01004), and the Scientific Research Project of the Hunan Provincial Health Commission (No.20254489).

doi: 10.46234/ccdcw2026.097

[#] Corresponding authors: Kaiwei Luo, 87616498@qq.com; Zeyu Zhao, zhaozeyu234@xmu.edu.cn.

[1] State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, Xiamen City, Fujian Province, China; [2] National Innovation Platform for Industry-Education Integration in Vaccine Research, Xiamen University, Xiamen City, Fujian Province, China; [3] Geographic Data Science Lab (GDSL), University of Liverpool, Liverpool, Merseyside, United Kingdom; [4] Geospatial, Analytics, Research & Insights (GARI) Team, UK Health Security Agency (UKHSA), London, United Kingdom; [5] CIRAD, Intertryp, Montpellier, France; [6] University Nice Côte d'Azur, Nice, France; [7] Department of Epidemiology, Biostatistics, Population Studies and Health Promotion, Faculty of Public Health, Universitas AIRLANGGA (UNAIR), Surabaya, Indonesia; [8] Department of Pathology, Faculty of Medicine-Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; [9] Department of Health, Faculty of Vocational Studies, Universitas Airlangga, Surabaya, Indonesia; [10] Hunan Provincial Center for Disease Control and Prevention, Changsha City, Hunan Province, China; [11]Aix-MarseilleUniversity,IHUMéditerranée-Infection,Marseille,France; [12] Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South, Changsha City, Hunan Province, China; [13] WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, UK.

[&] Joint first authors.

Submitted: January 19, 2026
Accepted: May 08, 2026
Issued: May 15, 2026

REFERENCES

1. Andrei G, Snoeck R. Differences in pathogenicity among the mpox virus clades: impact on drug discovery and vaccine development. Trends Pharmacol Sci 2023;44(10):719 – 39. https://doi.org/10.1016/j.tips.2023.08.003.
2. Walter K, Malani PN. Update on mpox. JAMA 2023;329(21):1892. https://doi.org/10.1001/jama.2023.9142.
3. Mauldin MR, McCollum AM, Nakazawa YJ, Mandra A, Whitehouse ER, Davidson W, et al. Exportation of monkeypox virus from the African continent. J Infect Dis 2022;225(8):1367 – 76. https://doi.org/10.1093/infdis/jiaa559.
4. Thornhill JP, Gandhi M, Orkin C. mpox: the reemergence of an old disease and inequities. Annu Rev Med 2024;75:159 – 75. https://doi.org/10.1146/annurev-med-080122-030714.
5. Precious ND, Agboola P, Oluwatimilehin O, Olakunle OK, Olaniyi P, Adiatu AI, et al. Re-emergence of monkeypox virus outbreak in Nigeria: epidemic preparedness and response (Review-Commentary). Ann Med Surg 2023;85(8):3990 – 6. https://doi.org/10.1097/MS9.0000000000001069.
6. Diaz Brochero C, Nocua-Báez LC, Cortes JA, Charniga K, Buitrago-Lopez A, Cucunubá ZM. Decoding mpox: a systematic review and meta-analysis of the transmission and severity parameters of the 2022-2023 global outbreak. BMJ Glob Health 2025;10(1):e016906. https://doi.org/10.1136/bmjgh-2024-016906.
7. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467 – 73. https://doi.org/10.7326/M18-0850.
8. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics—I. Bull Math Biol 1991;53(1-2):33 – 55. https://doi.org/10.1007/BF02464423.
9. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics—II. The problem of endemicity. Bull Math Biol 1991;53(1-2):57 – 87. https://doi.org/10.1007/BF02464424.
10. Case BKM, Young JG, Hébert-Dufresne L. Accurately summarizing an outbreak using epidemiological models takes time. R Soc Open Sci 2023;10(9):230634. https://doi.org/10.1098/rsos.230634.
11. Banuet-Martinez M, Yang Y, Jafari B, Kaur A, Butt ZA, Chen HH, et al. Monkeypox: a review of epidemiological modelling studies and how modelling has led to mechanistic insight. Epidemiol Infect 2023;151:e121. https://doi.org/10.1017/S0950268823000791.
12. Liao LC, Hsu CY, Chen HH, Lai CC. Estimating the global spread of epidemic human monkeypox with Bayesian directed acyclic graphic model. Vaccines (Basel) 2023;11(2):468. https://doi.org/10.3390/vaccines11020468.
13. Li ZH, Zhang L. Research on a stage-structured model of monkeypox transmission. Math Appl 2025;38(4):1110 – 21. https://doi.org/10.13642/j.cnki.42-1184/o1.2025.04.023.
14. Jin SH, Guan T, Endo A, Gan G, Janhavi A, Hu G, et al. Effectiveness of different border control strategies for reducing mpox importation risk: a modelling study. Lancet Reg Health Southeast Asia 2025;35:100565. https://doi.org/10.1016/j.lansea.2025.100565.
15. Tamakloe MD, Kuuwill A, Osumanu I, Siripi H. Mathematical modelling and time series clustering of mpox outbreak: a comparative study of the top 10 affected countries and implications for future outbreak management. Glob Epidemiol 2025;10:100214. https://doi.org/10.1016/j.gloepi.2025.100214.
16. Kaftan D, Kim HY, Ko C, Howard JS, Dalal P, Yamamoto N, et al. Performance analysis of mathematical methods used to forecast the 2022 New York City mpox outbreak. J Med Virol 2024;96(8):e29791. https://doi.org/10.1002/jmv.29791.
17. Nivetha S, Das P, Ghosh M. Threshold dynamics and epidemic-informed machine learning for forecasting of mpox: a U. S. case study. Chaos 2025;35(11):113118. https://doi.org/10.1063/5.0299032.
18. Rabiu M, Shatalov M, Malesela K. Bayesian forecasting of monkeypox epidemics in Central and East Africa using SEIR model and cumulative case data. Eur Phys J Plus 2025;140(12):1187. https://doi.org/10.1140/epjp/s13360-025-07126-x.
19. Pesantes-Grados P, Escalante-Ccoyllo N, Marín-Machuca O, Zambrano-Cabanillas AW, Ango-Aguilar H, Marín-Sánchez O, et al. A compartmental mathematical model to assess the impact of vaccination, isolation, and key epidemiological parameters on mpox control. Med Sci (Basel) 2025;13(4):226. https://doi.org/10.3390/medsci13040226.
20. Lin YC, Wen TH, Shih WL, Vermund SH, Fang CT. Impact of vaccination and high-risk group awareness on the mpox epidemic in the United States, 2022-2023: a modelling study. EClinicalMedicine 2024;68:102407. https://doi.org/10.1016/j.eclinm.2023.102407.
21. Qian M, Li DY, Hao ZH, Hu SJ, Li WD. An epidemiological model of monkeypox: model prediction and control application. BMC Infect Dis 2025;25(1):485. https://doi.org/10.1186/s12879-025-10873-y.
22. Alzubaidi AM, Othman HA, Ullah S, Ahmad N, Alam MM. Analysis of monkeypox viral infection with human to animal transmission via a fractional and fractal-fractional operators with power law kernel. Math Biosci Eng 2023;20(4):6666 – 90. https://doi.org/10.3934/mbe.2023287.
23. Majee S, Jana S, Kar TK. Dynamical analysis of monkeypox transmission incorporating optimal vaccination and treatment with cost-effectiveness. Chaos 2023;33(4):043103. https://doi.org/10.1063/5.0139157.
24. Khondaker F, Kamrujjaman M. Stochastic analysis of mpox epidemiology with vaccination strategies and environmental persistence. Sci Rep 2025;15(1):44369. https://doi.org/10.1038/s41598-025-28135-x.
25. Peter OJ, Babasola O, Ojo MM, Omame A. A mathematical model for assessing the effectiveness of vaccination in controlling mpox dynamics and mitigating disease burden in Nigeria and the Democratic Republic of Congo. J Appl Math Comput 2025;71(4):5729 – 56. https://doi.org/10.1007/s12190-025-02455-7.
26. Idisi IO, Oshinubi K, Sewanu VB, Yahaya MM, Olagbami OS, Edogbanya HO. Investigating mpox Strain dynamics using computational and data-driven approaches. Viruses 2025;17(2):154. https://doi.org/10.3390/v17020154.
27. Alharbi TD, Hasan R. Modeling monkeypox dynamics with human-rodent interactions and waning vaccination. AIMS Math 2025;10(8):18660 – 79. https://doi.org/10.3934/math.2025834.
28. Hassan AH, Aldila D, Aziz MHN, Kamalia PZ, Peter OJ. An optimal control problem for a monkeypox transmission model with activity-driven contact patterns. Braz J Phys 2025;55(5):213. https://doi.org/10.1007/s13538-025-01844-4.
29. Li S, Khan AA, Miao WL, Farhan M, Ullah S, AlQahtani SA, et al. A novel intelligent computing approach for modeling the population dynamics of monkeypox infection. Eur Phys J Plus 2025;140(5):441. https://doi.org/10.1140/epjp/s13360-025-06365-2.
30. Ngungu M, Addai E, Adeniji A, Adam UM, Oshinubi K. Mathematical epidemiological modeling and analysis of monkeypox dynamism with non-pharmaceutical intervention using real data from United Kingdom. Front Public Health 2023;11:1101436. https://doi.org/10.3389/fpubh.2023.1101436.
31. Khan MA, DarAssi MH, Ahmad I, Seyam NM, Alzahrani E. The transmission dynamics of an infectious disease model in fractional derivative with vaccination under real data. Comput Biol Med 2024;181:109069. https://doi.org/10.1016/j.compbiomed.2024.109069.
32. Hao YR, Zhao C. Stability of monkeypox virus infectious disease model with vaccine immunity and environmental infection. J Tianjin Norm Univ (Nat Sci Ed) 2025;45(5):9 – 15,21. https://doi.org/10.19638/j.issn1671-1114.20250502.
33. Agbata BC, Cenaj E, Dervishi R, Danjuma YJ, Shior MMA, Abah E, et al. Fractional-order mathematical model for monkeypox transmission dynamics using the Atangana-Baleanu Caputo operator. BMC Infect Dis 2025;25(1):1000. https://doi.org/10.1186/s12879-025-11383-7.
34. Onifade AA, Akindele OA, Ahmad I, Khan MA, Mat Isa N, Alzahrani E. Modeling monkeypox transmission with a compartmental framework to evaluate testing, isolation and public awareness strategies.

Sci Rep 2025;15(1):27236. https://doi.org/10.1038/s41598-025-10852-y.
35. Al Qurashi M, Rashid S, Alshehri AM, Jarad F, Safdar F. New numerical dynamics of the fractional monkeypox virus model transmission pertaining to nonsingular kernels. Math Biosci Eng 2023;20(1):402 – 36. https://doi.org/10.3934/mbe.2023019.
36. Wireko FA, Martey JN, Adu IK, Owusu BE, Ndogum S, Asamoah JKK. Modeling the impact of double-dose vaccination and saturated transmission dynamics on mpox control. Eng Rep 2025;7(5):e70144. https://doi.org/10.1002/eng2.70144.
37. Peter OJ, Oguntolu FA, Ojo MM, Oyeniyi AO, Jan R, Khan I. Fractional order mathematical model of monkeypox transmission dynamics. Phys Scr 2022;97(8):084005. https://doi.org/10.1088/1402-4896/ac7ebc.
38. El-Mesady A, Elsonbaty A, Adel W. On nonlinear dynamics of a fractional order monkeypox virus model. Chaos Solitons Fractals 2022;164:112716. https://doi.org/10.1016/j.chaos.2022.112716.
39. Spath T, Brunner-Ziegler S, Stamm T, Thalhammer F, Kundi M, Purkhauser K, et al. Modeling the protective effect of previous compulsory smallpox vaccination against human monkeypox infection: from hypothesis to a worst-case scenario. Int J Infect Dis 2022;124:107 – 12. https://doi.org/10.1016/j.ijid.2022.09.022.
40. Yapışkan D, Yurtoğlu M, Avcı D, İskender Eroğlu BB, Bonyah E. A novel model for monkeypox disease: system analysis and optimal preventive strategies. Iran J Sci 2023;47(5):1665 – 77. https://doi.org/10.1007/s40995-023-01525-4.
41. Fernandes GD, Maldonado V. Behavioral aspects and the transmission of monkeypox: a novel approach to determine the probability of transmission for sexually transmissible diseases. Infect Dis Model 2023;8(3):842 – 54. https://doi.org/10.1016/j.idm.2023.07.008.
42. Yuan P, Tan Y, Yang L, Aruffo E, Ogden NH, Bélair J, et al. Assessing transmission risks and control strategy for monkeypox as an emerging zoonosis in a metropolitan area. J Med Virol 2023;95(1):e28137. https://doi.org/10.1002/jmv.28137.
43. Liu BT, Farid S, Ullah S, Altanji M, Nawaz R, Wondimagegnhu Teklu S. Mathematical assessment of monkeypox disease with the impact of vaccination using a fractional epidemiological modeling approach. Sci Rep 2023;13(1):13550. https://doi.org/10.1038/s41598-023-40745-x.
44. Collins OC, Duffy KJ. Dynamics and control of mpox disease using two modelling approaches. Model Earth Syst Environ 2024;10(2):1657 – 69. https://doi.org/10.1007/s40808-023-01862-8.
45. Peter OJ, Babasola O, Ojo MM, Omame A. Modelling the transmission of mpox with a case study in Nigeria and the Democratic Republic of the Congo (DRC). Comput Methods Differ Equations 2025;13(3):995 – 1011. https://doi.org/10.22034/cmde.2024.62086.2711.
46. Khan U, Ali N, Ahmad I, Khan H, Almutairi DK, Alzabut J, et al. Monkeypox dynamical system with stability and computational analysis of the transmission. J Math Comput Sci 2025;39(3):300 – 24. https://doi.org/10.22436/jmcs.039.03.01.
47. Awoke TD, Kassa SM, Terefe YA, Asfaw MD. Modeling on cost-effectiveness of monkeypox disease control strategies with consideration of environmental transmission effects in the presence of vaccination. Model Earth Syst Environ 2024;10(5):6105 – 32. https://doi.org/10.1007/s40808-024-02108-x.
48. Al-Shomrani MM, Musa SS, Yusuf A. Unfolding the transmission dynamics of monkeypox virus: an epidemiological modelling analysis. Mathematics 2023;11(5):1121. https://doi.org/10.3390/math11051121.
49. Mrope FM, Nyerere N, Iddi AJ, Karugila GK. Modeling the impact of traditional remedies on the transmission dynamics of mpox. Sci Afr 2025;30:e03085. https://doi.org/10.1016/j.sciaf.2025.e03085.
50. Yuan P, Tan Y, Yang L, Aruffo E, Ogden NH, Bélair J, et al. Modeling vaccination and control strategies for outbreaks of monkeypox at gatherings. Front Public Health 2022;10:1026489. https://doi.org/10.3389/fpubh.2022.1026489.
51. Gao J, Zhou C, Liang HW, Jiao R, Wheelock ÅM, Jiao K, et al. Monkeypox outbreaks in the context of the COVID-19 pandemic: network and clustering analyses of global risks and modified SEIR prediction of epidemic trends. Front Public Health 2023;11:1052946. https://doi.org/10.3389/fpubh.2023.1052946.
52. Xiong ZX, Xue L, Li XZ, Zhang YF. Assessing vaccine strategies for mpox outbreak in New York City using an age-structure model. BMC Infect Dis 2024;24(1):1078. https://doi.org/10.1186/s12879-024-09551-2.
53. Li T, Guo XH, Wang XL, Chen TM. Temporal and age-structured analysis of mpox spread in the 2022 Global outbreak: data-assimilation insights for epidemic control. Infect Dis Poverty 2025;14(1):100. https://doi.org/10.1186/s40249-025-01369-7.
54. Omame A, Han Q, Iyaniwura SA, Ebenezer A, Bragazzi NL, Wang XY, et al. Understanding the impact of HIV on mpox transmission in the MSM population: a mathematical modeling study. Infect Dis Model 2024;9(4):1117 – 37. https://doi.org/10.1016/j.idm.2024.05.008.
55. Xiridou M, van Wees DA, Adam P, Miura F, Op de Coul E, Reitsema M, et al. Combining mpox vaccination and behavioural changes to control possible future mpox resurgence among men who have sex with men: a mathematical modelling study. BMJ Public Health 2025;3(2):e002682. https://doi.org/10.1136/bmjph-2025-002682.
56. Bragazzi NL, Han Q, Iyaniwura SA, Omame A, Shausan A, Wang XY, et al. Adaptive changes in sexual behavior in the high-risk population in response to human monkeypox transmission in Canada can help control the outbreak: Insights from a two-group, two-route epidemic model. J Med Virol 2023;95(4):e28575. https://doi.org/10.1002/jmv.28575.
57. Yang ST, Guo XH, Zhao ZY, Abudunaibi B, Zhao YK, Rui J, et al. Possibility of mpox viral transmission and control from high-risk to the general population: a modeling study. BMC Infect Dis 2023;23(1):119. https://doi.org/10.1186/s12879-023-08083-5.
58. Omame A, Iyaniwura SA, Ebenezer A, Han Q, Wang XY, Bragazzi NL, et al. Pre-exposure vaccination in the high-risk population is crucial in controlling mpox resurgence in Canada. BMC Infect Dis 2026;26(1):14. https://doi.org/10.1186/s12879-025-12172-y.
59. Garakani S, Flores L, Alvarez-Pardo G, Rychtář J, Taylor D. The effect of heterogeneity of relative vaccine costs on the mean population vaccination rate with mpox as an example. J Theor Biol 2025;602-603:112062. http://dx.doi.org/10.1016/j.jtbi.2025.112062.
60. Knight J, Tan DHS, Mishra S. Maximizing the impact of limited vaccine supply under different early epidemic conditions: a 2-city modelling analysis of monkeypox virus transmission among men who have sex with men. CMAJ 2022;194(46):E1560 – 7. https://doi.org/10.1503/cmaj.221232.
61. Huang QR, Sun YX, Jia MM, Jiang MY, Xu YS, Feng LZ, et al. An effectiveness study of vaccination and quarantine combination strategies for containing mpox transmission on simulated college campuses. Infect Dis Model 2024;9(3):805 – 15. https://doi.org/10.1016/j.idm.2024.04.004.
62. Zhang L, Huang JP, Yan W, Zhao YJ, Wang DF, Chen B. Global prediction for mpox epidemic. Environ Res 2024;243:117748. https://doi.org/10.1016/j.envres.2023.117748.
63. Cao YW, Fang WB, Chen YY, Zhang HC, Ni RY, Pan GX. Simulating the impact of optimized prevention and control measures on the transmission of monkeypox in the United States: a model-based study. J Med Virol 2024;96(2):e29419. https://doi.org/10.1002/jmv.29419.
64. Ochieng FO. Mathematical modeling of mpox virus dynamics with a case study of Africa. Model. Earth Syst Environ 2025;11:193. https://doi.org/10.1007/s40808-025-02369-0.
65. Rabiu M, Fagbemigun B, Fadugba S, Shatalov M, Malesela K, Adeniji A. Quantifying mpox transmission and control: a regional analysis of vaccination strategies in East Africa. Infect Dis Model 2026;11(1):29 – 46. https://doi.org/10.1016/j.idm.2025.09.001.
66. Savinkina A, Kindrachuk J, Bogoch II, Rimoin AW, Hoff NA, Shaw SY, et al. Modelling vaccination approaches for mpox containment and mitigation in the Democratic Republic of the Congo. Lancet Glob Health 2024;12(12):e1936 – 44. https://doi.org/10.1016/S2214-109X(24)00384-X.
67. Zhang XS, Niyomsri S, Mandal S, Mohammed H, Mindlin M,

Dugbazah B, et al. Cost-effectiveness of vaccination strategies to control future mpox outbreaks in England: a modelling study. Lancet Reg Health Eur 2025;55:101364. https://doi.org/10.1016/j.lanepe.2025.101364.
68. Xiridou M, Miura F, Adam P, Op de Coul E, de Wit J, Wallinga J. The fading of the mpox outbreak among men who have sex with men: a mathematical modelling study. J Infect Dis 2024;230(1):e121 – 30. https://doi.org/10.1093/infdis/jiad414.
69. Jing Y, Hou Q. Reproduction number estimation and dynamic analysis of mpox transmission in China. J Shandong Univ (Nat Sci) 2025;60(4):118 – 26. https://doi.org/10.6040/j.issn.1671-9352.0.2023.509.
70. Bragazzi NL, Iyaniwura SA, Han Q, Woldegerima WA, Kong JD. Quantifying the basic reproduction number and underestimated fraction of mpox cases worldwide at the onset of the outbreak. J R Soc Interface 2024;21(216):20230637. https://doi.org/10.1098/rsif.2023.0637.
71. Alqahtani RT, Musa SS, Inc M. Modeling the role of public health intervention measures in halting the transmission of monkeypox virus. AIMS Math 2023;8(6):14142 – 66. https://doi.org/10.3934/math.2023723.
72. Al-Raeei M. The study of human monkeypox disease in 2022 using the epidemic models: herd immunity and the basic reproduction number case. Ann Med Surg (Lond) 2023;85(2):316 – 21. https://doi.org/10.1097/MS9.0000000000000229.
73. Puebla J. Analysis of monkeypox transmission dynamics incorporating quarantine and surface contamination. Eur J Pure Appl Math 2025;18(4):6750. https://doi.org/10.29020/nybg.ejpam.v18i4.6750.
74. Acosta-Tovar A, Lopes F. Who should be controlled? The role of asymptomatic individuals, isolation and switching in the dominant transmission route in classical and network epidemic models. Bull Math Biol 2025;87(10):143. https://doi.org/10.1007/s11538-025-01521-2.
75. Acheneje GO, Omale D, Atokolo W, Celestine AB, Abah E, Dervishi R, et al. Mathematical modelling on the transmission dynamics of the co-infection of COVID-19 and Monkeypox with treatment as a control strategy. Discov Public Health 2025;22(1):535. https://doi.org/10.1186/s12982-025-00893-z.
76. Ouemba Tassé AJ, Terefe Y, Lubuma J. Assessing the influence of HIV on the spread of mpox disease. Math Biosci 2025;387:109499. https://doi.org/10.1016/j.mbs.2025.109499.
77. Omame A, Iyaniwura SA, Han Q, Ebenezer A, Bragazzi NL, Wang XY, et al. Dynamics of mpox in an HIV endemic community: a mathematical modelling approach. Math Biosci Eng 2025;22(2):225 – 59. https://doi.org/10.3934/mbe.2025010.
78. Zhang ZH, Chen Y. Transmission dynamics of mpox virus and HIV co-infection and its optimal control in MSM population. Adv Cont Discr Mod 2025;2025(1):169. https://doi.org/10.1186/s13662-025-04029-0.
79. Fatahillah HA, Chávez JP, Aldila D. Emergence of chaos in a coinfection model of HIV/AIDS and mpox with treatment constraints. Nonlinear Dyn 2025;113(23):33005 – 33. https://doi.org/10.1007/s11071-025-11803-y.
80. Liu BY, Jing SL, Xiang H, Huo HF. Modeling the transmission dynamic of mpox virus with vertical transmission. Math Methods Appl Sci 2025;48(14):13517 – 29. https://doi.org/10.1002/mma.11119.
81. Rabiu M, Dansu EJ, Mogbojuri OA, Idisi IO, Yahaya MM, Chiwira P, et al. Modeling the sexual transmission dynamics of mpox in the United States of America. Eur Phys J Plus 2024;139(3):250. https://doi.org/10.1140/epjp/s13360-024-05020-6.
82. Xue L, Zhou S, Xiong ZX, Zhang YX. Modeling and assessing the risk of mpox outbreak in China through international travel. J Math Anal Appl 2026;553(1):129837. https://doi.org/10.1016/j.jmaa.2025.129837.
83. Rabiu S, Ali MKM. Modeling the transmission dynamics of mpox virus. Sci Rep 2025;15(1):34876. https://doi.org/10.1038/s41598-025-16303-y.
84. McCollum AM, Damon IK. Human monkeypox. Clin Infect Dis 2014;58(2):260 – 7. https://doi.org/10.1093/cid/cit703.
85. Xiu FY, Doyle CM, Anato JLF, Knight J, Wang LW, Cox J, et al. Impact of interventions on mpox transmission during the 2022 outbreak in Canada: a mathematical modeling study of three different cities. Int J Infect Dis 2025;153:107792. https://doi.org/10.1016/j.ijid.2025.107792.
86. Machado-Marques S, Romero-Leiton JP, Nasri B, Moyles IR. Considering the effects of pair formation dynamics on mpox and HIV co-infection in the gbMSM community. Sci Rep 2025;15(1):36028. https://doi.org/10.1038/s41598-025-20039-0.
87. Obiajulu EF, Iheonu NO, Araka NN, Omame A. Stability and bifurcation analysis in a co-dynamical model for mpox and syphilis incorporating intervention measures using real data from USA. Model Earth Syst Environ 2026;12(1):59. https://doi.org/10.1007/s40808-025-02666-8.
88. Americo JL, Earl PL, Moss B. Virulence differences of mpox (monkeypox) virus clades I, IIa, and IIb. 1 in a small animal model. Proc Natl Acad Sci USA 2023;120(8):e2220415120. https://doi.org/10.1073/pnas.2220415120.
89. Wawina-Bokalanga T, Merritt S, Kinganda-Lusamaki E, Jansen D, Halbrook M, O'Toole Á, et al. Epidemiology and phylogenomic characterisation of two distinct mpox outbreaks in Kinshasa, DR Congo, involving a new subclade Ia lineage: a retrospective, observational study. Lancet 2025;406(10498):63 – 75. https://doi.org/10.1016/S0140-6736(25)00294-6.
90. Murayama H, Asakura TR, Dickens BL, Boyle D, Foo JH, Jin SH, et al. Roles of community and sexual contacts as drivers of clade I mpox outbreaks. medRxiv http://dx.doi.org/10.1101/2024.10.15.24315554.
91. 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.
92. 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.
93. Ahmed ZL, Islam R. The emergence of a novel mpox virus strain (clade Ib) in Central Africa: a global public health concern. J Infect Public Health 2025;18(7):102781. https://doi.org/10.1016/j.jiph.2025.102781.
94. Soto-Cabezas MG, Vásquez-Mejía A, Calderón M, Reyes-Vega M, Blancas-Blas YI, Soriano-Moreno AN, et al. Transmission dynamics of mpox in Peru during the 2022 outbreak. Am J Trop Med Hyg 2025;113(2):387 – 94. https://doi.org/10.4269/ajtmh.24-0361.
95. World Health Organization. Third meeting of the International Health Regulations (2005) Emergency Committee regarding the upsurge of mpox 2024. 2025. https://www.who.int/news/item/17-03-2025-third-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-upsurge-of-mpox-2024. [2026-2-2].
96. 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.

Vital Surveillances

Spatiotemporal Transmission Characteristics of Global Monkeypox — Worldwide, January 2022–September 2025

Wenqi Yin[1,&]; Fu Li[1,2,&]; Mingchen Zhao[3]; Xuefeng Huang[4]; Qiangru Huang[4]; Mengyuan Wang[3]; Zhuowei Luo[3]; Yonghong Liu[3]; Mengyao Li[3]; Da Huo[1,2]; Zhongjie Li[4,5,6,7]; Xiaoli Wang[1,3,5,#]

ABSTRACT

Introduction: Since 2022, the global monkeypox (mpox) epidemic has undergone clade replacement from clade IIb to clade Ib, accompanied by a geographic shift in the epidemic's epicenter. However, systematic analyses spanning the entire epidemic cycle from 2022 to 2025 remain limited. Integrating global epidemiological and genomic data is essential for determining spatiotemporal transmission dynamics.

Methods: Global mpox surveillance data, mpox virus whole-genome sequences, and demographic data were collected. Retrospective spatiotemporal scan statistics were used to identify significant transmission clusters. Phylogenetic trees were used to characterize evolutionary relationships among clades, and haplotype network analysis was used to infer transmission pathways.

Results: A total of 165,244 mpox cases were reported globally, showing a two-phase pattern. From January 2022 to August 2023, clade IIb predominated, accounting for 96.18% of cases in Europe and the Americas. From September 2023 to September 2025, the epicenter shifted to Africa, where clade Ib became dominant; during this phase, Africa accounted for 71.48% of global cases. Spatiotemporal scan analysis identified 10 significant clusters, with the highest relative risk observed in the West African cluster in 2025. Phylogenetic analyses linked clade IIb to the European and American phases, whereas clade Ib was associated with the African phase. The haplotype network suggested that Africa is the core transmission region of clade Ib, with Europe serving as a dissemination hub.

Conclusion: Mpox evolved from IIb-driven global dissemination to Ib-dominated regional clustering in Africa, with clade replacement driving the geographic shifts. The current risk is concentrated in Africa, underscoring the need for strengthened surveillance, genomic tracing, equitable access to resources, and targeted interventions.

Since May 2022, the mpox virus (MPXV) clade IIb, originating in Africa, has triggered a global outbreak affecting more than 100 countries. This outbreak marked a fundamental shift in mpox from a regionally endemic zoonotic disease to a sustained global public health threat characterized by human-to-human transmission (*1*). The World Health Organization (WHO) declared it a Public Health Emergency of International Concern (PHEIC) in July 2022 and terminated this status in May 2023. However, from late 2023 to 2024, the renewed transmission and international spillover of clade Ib, originating in Africa, prompted the WHO to redeclare a PHEIC in August 2024. Although this status was lifted in September 2025, the ongoing global transmission risk continues to be emphasized (*2*).

Although extensive studies have examined the global outbreak of MPXV driven by clade IIb (*3*), and the subsequent African resurgence associated with clade Ib (*4–5*), a systematic analysis of the complete epidemic trajectory from 2022 to 2025 remains lacking. Previous studies have largely focused on individual clades, specific regions, or limited time periods (*6–7*) and have failed to capture the macro-level intercontinental shift in epidemic centers driven by alternating clade dominance, as well as the micro-level mechanisms of viral adaptive evolution.

This study aimed to conduct an integrated analysis of the global mpox epidemic from January 2022 to September 2025 by combining macroepidemiological evidence (spatiotemporal distribution of cases) with microvirological insights (genomic evolutionary characteristics). The findings provide a robust scientific basis for global surveillance systems and differentiated prevention and control strategies.

METHODS

Data Sources

Confirmed global mpox case data were obtained from the WHO website. Viral clade distribution and whole-genome sequence data were retrieved from the Global Initiative on Sharing All Influenza Data (GISAID) database. The population data for each country were sourced from the World Bank.

Analytical Methods

Spatial distribution maps were generated using ArcGIS 10.8.1 (Environmental Systems Research Institute, Inc., Redlands, California, USA). Bar charts and heatmaps were produced using the "ggplot2" and "pheatmap" packages in R (version 4.4.2, R Foundation for Statistical Computing, Vienna, Austria). A retrospective spatiotemporal scan analysis was conducted using SaTScan v10.2 (Information Management Services, Inc., Rockville, Maryland, USA) under a Poisson model, with a maximum spatial window of 30% of the at-risk population, a temporal window of 6–12 months, and a minimum cluster size of 1,000 cases. Cluster strength was assessed using the log-likelihood ratio, observed-to-expected ratio, and relative risk (*RR*), with $P<0.05$ considered statistically significant.

The eligible MPXV whole-genome sequences were filtered using GISAID (GISAID Initiative, Munich, Bavaria, Germany). Multiple sequence alignments were performed using the Multiple Alignment using Fast Fourier Transform (MAFFT) software (Osaka University, Suita, Osaka, Japan). Phylogenetic trees were constructed using IQ-TREE 2.0 (University of Vienna, Vienna, Austria) with the maximum-likelihood method and 1,000 bootstrap replicates, and visualized using iTOL v7 (biobyte solutions GmbH, Heidelberg, Baden-Württemberg, Germany). Quality-controlled whole-genome sequences from clade Ib were used to construct haplotype networks with the modified Templeton-Crandall-Singh (TCS) algorithm in fast Haplotype Networker (fastHaN), and visualization was performed with tcsBU (University of Porto, Porto, Portugal).

RESULTS

Epidemiological Overview of Global Monkeypox

Between January 2022 and September 2025, 165,244 mpox cases were reported across 135 countries. Regarding clade distribution (Table 1), clade IIb was predominantly detected in Europe and the Americas, whereas clades Ia, Ib, and IIa were mainly concentrated in Africa, with sporadic detections in other regions.

The global mpox epidemic exhibited a pronounced two-phase pattern (Supplementary Figure S1 and Supplementary Table S1, available at https://weekly.chinacdc.cn/). During the first phase (January 2022–August 2023), clade IIb predominated, with rapid expansion beginning in May 2022 and peaking in August 2022. In total, 96.18% of cases occurred in Europe and the Americas. In the second phase (September 2023 to September 2025), the epidemic center shifted to Africa, where clade Ib became the dominant strain and a new peak emerged in May 2025. During this phase, Africa accounted for 71.48% of global cases, while the Western Pacific, Eastern Mediterranean, and South-East Asian regions reported persistently low incidence levels.

Country-level heatmaps (Figure 1) show that Europe and the Americas experienced concentrated mpox outbreaks between June and August 2022, with Spain (6.02/100,000 in July) and Luxembourg (4.43/100,000 in August) reporting the highest monthly mpox incidence. From September 2023 onward, mpox incidence markedly increased in the Democratic Republic of the Congo (DRC), followed by Burundi and Uganda, with increases after August 2024. In early 2025, a new cluster emerged in West Africa, and Sierra Leone reached a peak incidence of 26.37/100,000 in May, the highest globally.

Spatiotemporal Clustering Analysis

A retrospective spatiotemporal scan analysis identified 10 significant transmission clusters with substantial heterogeneity in their spatial and temporal distributions. The five highest-risk clusters (Table 2), in chronological order, were as follows: a European cluster involving 19 countries from May to October 2022 (centered in the United Kingdom); a North American cluster from June to November 2022 (centered in the United States); a South American cluster involving nine countries from July to December 2022 (centered in Bolivia); a Central and Eastern African cluster from September 2024 to August 2025 (centered in the Democratic Republic of the Congo); and a West African cluster comprising Guinea, Sierra Leone, and Liberia from April to September 2025, which showed the highest *RR* (*RR*=86.69, $P<0.001$).

TABLE 1. Global distribution of mpox virus clades from January 2022 to September 2025.

MPXV clade	Number of countries	Countries	WHO region distribution (*n*, %)
Ia	15	Benin, Cameroon, Central African Republic, China, Democratic Republic of the Congo, Denmark, Gabon, Ireland, Liberia, Netherlands, Republic of the Congo, Sierra Leone, Sudan, Turkey, USA	AFRO: 8 (53.33); AMRO: 1 (6.67); EMRO: 1 (6.67); EURO: 4 (26.67); WPRO: 1 (6.67)
Ib	34	Angola, Australia, Belgium, Brazil, Burundi, Canada, China, Democratic Republic of the Congo, Ethiopia, France, Germany, India, Ireland, Italy, Japan, Kenya, Malawi, Oman, Pakistan, Qatar, Republic of the Congo, Rwanda, South Africa, South Sudan, Sweden, Switzerland, Tanzania, Thailand, Uganda, United Arab Emirates, UK, USA, Zambia, Zimbabwe	AFRO: 14 (41.18); AMRO: 3 (8.82); EMRO: 4 (11.76); EURO: 8 (23.53); SEARO: 2 (5.88); WPRO: 3 (8.82)
IIa	9	Cote d'Ivoire, Denmark, France, Ghana, Guinea, Liberia, Netherlands, Sierra Leone, USA	AFRO: 5 (55.56); AMRO: 1 (11.11); EURO: 3 (33.33)
IIb	72	Argentina, Australia, Austria, Belgium, Benin, Brazil, Cambodia, Cameroon, Canada, Central African Republic, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Czech Republic, Democratic Republic of the Congo, Denmark, Dominican Republic, Ecuador, Egypt, Ethiopia, Finland, France, Georgia, Germany, Ghana, Guinea, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Lithuania, Luxembourg, Mexico, Nepal, Netherlands, Nigeria, North Macedonia, Oman, Pakistan, Paraguay, Peru, Philippines, Poland, Portugal, Republic of the Congo, Romania, Russia, Rwanda, Republic of Korea, Sierra Leone, Singapore, Slovakia, Slovenia, South Africa, Spain, Sudan, Sweden, Switzerland, Thailand, Togo, Turkey, UK, USA, Venezuela, Vietnam	AFRO: 14 (19.44); AMRO: 13 (18.06); EMRO: 5 (6.94); EURO: 28 (38.89); SEARO: 4 (5.56); WPRO: 8 (11.11)

Note: MPXV clade information was obtained from sequences available in the GISAID database as of September 30, 2025. n denotes the number of countries in each WHO region where a given MPXV clade was identified. The percentage (%) was calculated as the number of countries with that clade in each WHO region divided by the total number of countries with that clade across all WHO regions.
Abbreviation: MPXV=mpox virus; WHO=World Health Organization; SEARO=South-East Asia Region; EMRO=Eastern Mediterranean Region; EURO=European Region; AMRO=Region of the Americas; WPRO=Western Pacific Region; AFRO=African Region; USA=United States of America; UK=United Kingdom.

Phylogenetic and Haplotype Analyses

Phylogenetic analysis exhibited strong concordance between the staged evolution of the global mpox epidemic and the evolutionary dynamics of distinct viral clades (Figure 2A). Clade IIb showed extensive intercontinental dissemination and high genetic diversity in 2022, followed by a gradual decline after 2023, consistent with the epidemic phase dominated by Europe and the Americas. Conversely, clade Ib showed a highly localized transmission core in Africa, particularly in the Democratic Republic of the Congo and Uganda, with limited genetic diversity, reflecting the Africa-centered second phase of the epidemic. Clades Ia and IIa displayed more restricted transmission ranges than the other clades, indicating substantial differences in transmission adaptability among the MPXV clades.

Haplotype network analysis of clade Ib further showed that the DRC was the persistent core of regional transmission, while Uganda experienced short-term, high-intensity outbreaks. Other countries primarily showed sporadic importations or limited local transmission chains that were rapidly interrupted, indicating that clade Ib remained largely Africa-centered with limited global dissemination capacity (Figure 2B).

DISCUSSION

This study represents the first comprehensive integration of macroepidemiological and microvirological evidence to characterize the complete global mpox epidemic trajectory from January 2022 to September 2025. The results reveal a clear two-phase spatiotemporal pattern. We determined the intrinsic linkage between viral clade replacement (IIb to Ib) and intercontinental shifts in epidemic epicenters and addressed key gaps in previous studies that focused on single clades or localized spatiotemporal scales.

During the first phase, clade IIb drove a rapid expansion in Europe and the Americas, which was closely associated with dense social networks among men who have sex with men and high levels of international mobility (*8*–*10*). In the second phase, the epidemic center shifted to Africa, where clade Ib predominated and showed strong regional clustering. This pattern likely reflects Africa's role as a traditional reservoir of clade I mpox, which is characterized by

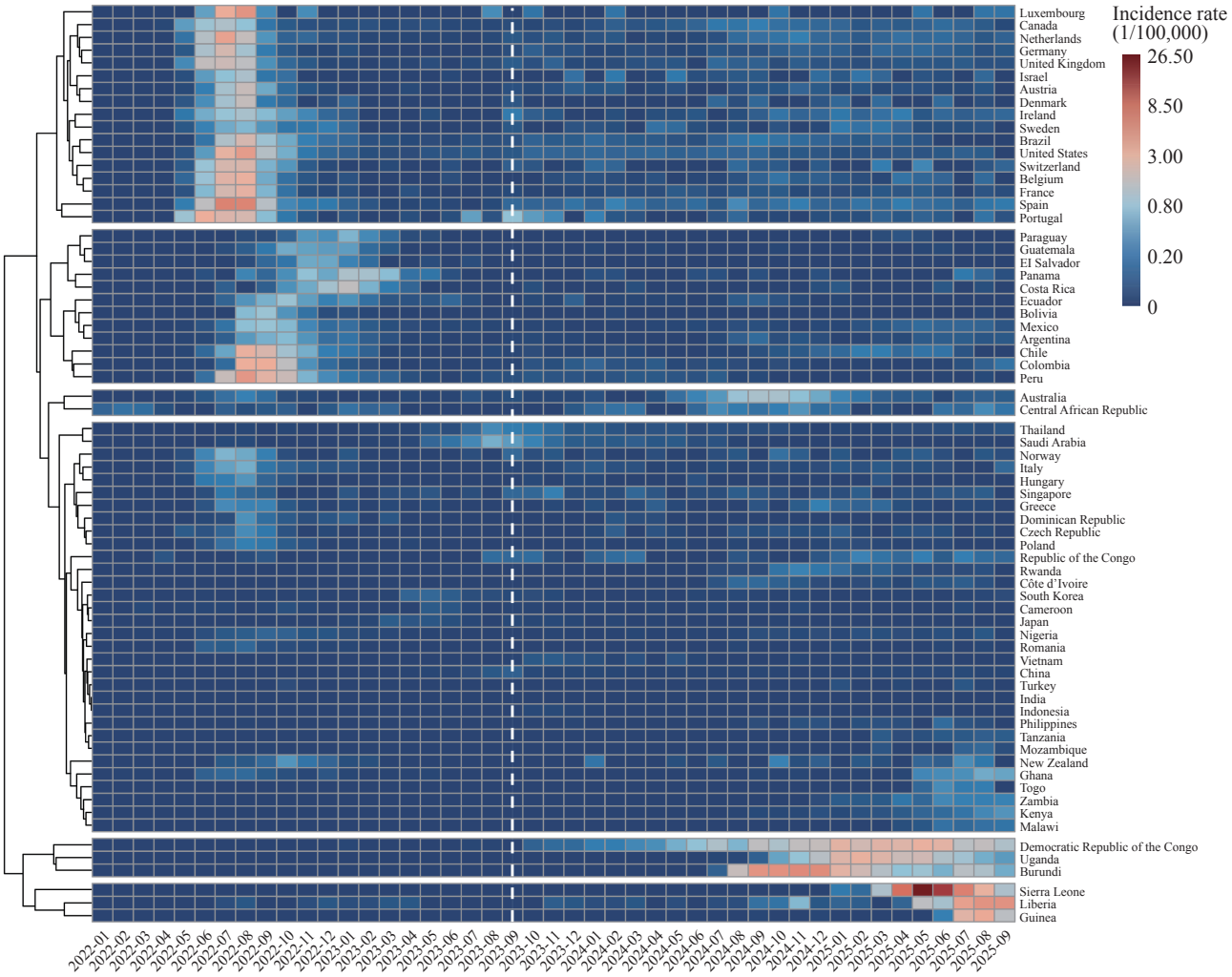


FIGURE 1. Heatmap of monthly mpox incidence rates among countries that reported more than 50 cumulative cases from January 2022 to September 2025.
Note: Countries with more than 50 cumulative cases were included, covering more than 99% of all reported mpox cases globally. Monthly incidence rates were calculated using the formula $\ln(x\times10^5+0.1)$. Hierarchical clustering of rows and columns was performed using the complete linkage method with Euclidean distances. A total of 69 countries were included, with rows representing countries and columns representing months. Color gradients indicate the magnitude of monthly incidence rates, with red indicating higher incidence and blue indicating lower incidence.

sustained zoonotic transmission cycles (*11*). Moreover, clade Ib transmission is primarily associated with close heterosexual contact (*12*), involves geographically constrained social networks, and emerges after widespread clade IIb circulation, when partial population immunity from prior infection or vaccination may have limited its interregional spread (*13*). The lack of widespread intercontinental dissemination of clades Ia and IIa further underscores the marked differences in transmission adaptability among MPXV clades. Notably, the West African cluster identified in 2025 exhibited the highest RR. This finding suggests that this region has become a current transmission hotspot, potentially driven by limited public health resources and insufficient surveillance capacity (*14–15*).

This study had several limitations. First, although 14,089 MPXV whole-genome sequences were retrieved from the GISAID during the study period, these genomes represent only a small proportion of the global mpox burden (165,244 cases across 135 countries from January 2022 to September 2025) (Supplementary Table S3, available at https://weekly.chinacdc.cn/). Therefore, the genomic dataset may not fully represent the global pandemic. This constraint likely reflects pronounced heterogeneity across countries and regions in sampling intensity, sequencing capacity, surveillance practices, and the timeliness of

TABLE 2. Characteristics of the five highest-risk mpox transmission clusters worldwide from January 2022 to September 2025.

Cluster	Time frame	Locations included	Number of cases	Observed/ expected	Relative risk	Log likelihood ratio	*P*
1	2022/5/1 to 2022/10/31	United Kingdom (cluster center), Ireland, Netherlands, Belgium, Luxembourg, France, Denmark, Germany, Switzerland, Andorra, Iceland, Spain, Portugal, Monaco, Czech Republic, Austria, Sweden, Slovenia, San Marino	23,409	20.98	24.28	50,544.05	<0.001
2	2022/6/1 to 2022/11/30	USA (cluster center), Canada, Mexico	34,641	22.47	28.16	78,303.17	<0.001
3	2022/7/1 to 2022/12/31	Bolivia (Plurinational State of) (cluster center), Paraguay, Peru, Brazil, Uruguay, Argentina, Chile, Ecuador, Colombia	21,460	16.98	19.36	41,871.78	<0.001
4	2024/9/1 to 2025/8/31	Democratic Republic of the Congo (cluster center), Rwanda, Burundi, Uganda	36,774	34.56	44.16	98,765.45	<0.001
5	2025/4/1 to 2025/9/30	Guinea (cluster center), Sierra Leone, Liberia	7,111	83.00	86.69	24,548.58	<0.001

sequence submission. Such spatiotemporal sampling imbalance may introduce ascertainment bias, thereby affecting regional comparisons and the interpretation of temporal trends. Specifically, given the widespread global circulation of clade IIb and the relatively high number of well-assembled sequences submitted from high-burden regions, the dataset likely provides a comprehensive representation of its genetic diversity. However, the limited sequencing capacity in Africa, further compounded by the exclusion of low-coverage sequences, may mean that the true diversity and extent of clade Ib transmission are not fully captured. These imbalances underscore the urgent need to enhance equity and balance in global genomic surveillance. Second, the inference of spatiotemporal transmission was limited by imperfect temporal and geographic concordance between epidemiological investigations and available genomic data. Nevertheless, the principal evolutionary patterns identified in this study are broadly consistent with the observed phase-specific macroepidemiological features. Specifically, clade IIb showed extensive intercontinental dissemination and relatively high genetic diversity by 2022, aligning with the Europe- and Americas-dominated phases, followed by a gradual decline after 2023. In contrast, clade Ib exhibited a more localized transmission focus in Africa with comparatively limited diversity, which is consistent with the subsequent Africa-centered phase. Taken together, this concordance supports the robustness of the core conclusions; however, inferences for regions with sparse genomic sampling should be interpreted with caution.

Currently, clade Ib continues to circulate in Africa, with sporadic importations and limited local transmission. In 2025, an additional outbreak driven by clade IIb occurred in Sierra Leone, indicating that the mpox epidemic caused by clade IIb persists. The G.1 lineage (A.2.2.1) detected in Sierra Leone carries a high proportion of apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3)-type mutations, suggesting potential human adaptation (*16*). However, whether this lineage has acquired enhanced transmissibility remains to be determined. Future mpox prevention and control efforts should adopt a framework of global collaboration; prioritize Africa, particularly West Africa, for enhanced surveillance, genomic sequencing, and molecular tracing; and promote equitable access to vaccines and antiviral therapies. Targeted interventions tailored to distinct transmission patterns and high-risk populations are essential to prevent the recurrence of global pandemics.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the Special Research Project for Capital Health Development (2024-2G-3016), the National Key Research and Development Program of China (2023YFC2308700), Prevention and Control of Emerging and Major Infectious Diseases–National Science and Technology Major Project (2025ZD01900800), Capital's Funds for Health Improvement and Research (2026-1G-4291), Key Research Projects of the Beijing Natural Science Foundation–Haidian District Joint Fund (L242053).

doi: 10.46234/ccdcw2026.098

[#] Corresponding author: Xiaoli Wang, wangxiaoli@bjcdc.org.

[1] School of Public Health, Capital Medical University, Beijing, China; [2] Institute for Infectious Disease and Endemic Disease Control, Beijing Center for Disease Prevention and Control, Beijing, China; [3] Beijing Office of Global Health, Beijing Center for Disease Prevention and Control, Beijing, China; [4] School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [5] Public Health Emergency Management Innovation Center of Beijing Higher Education

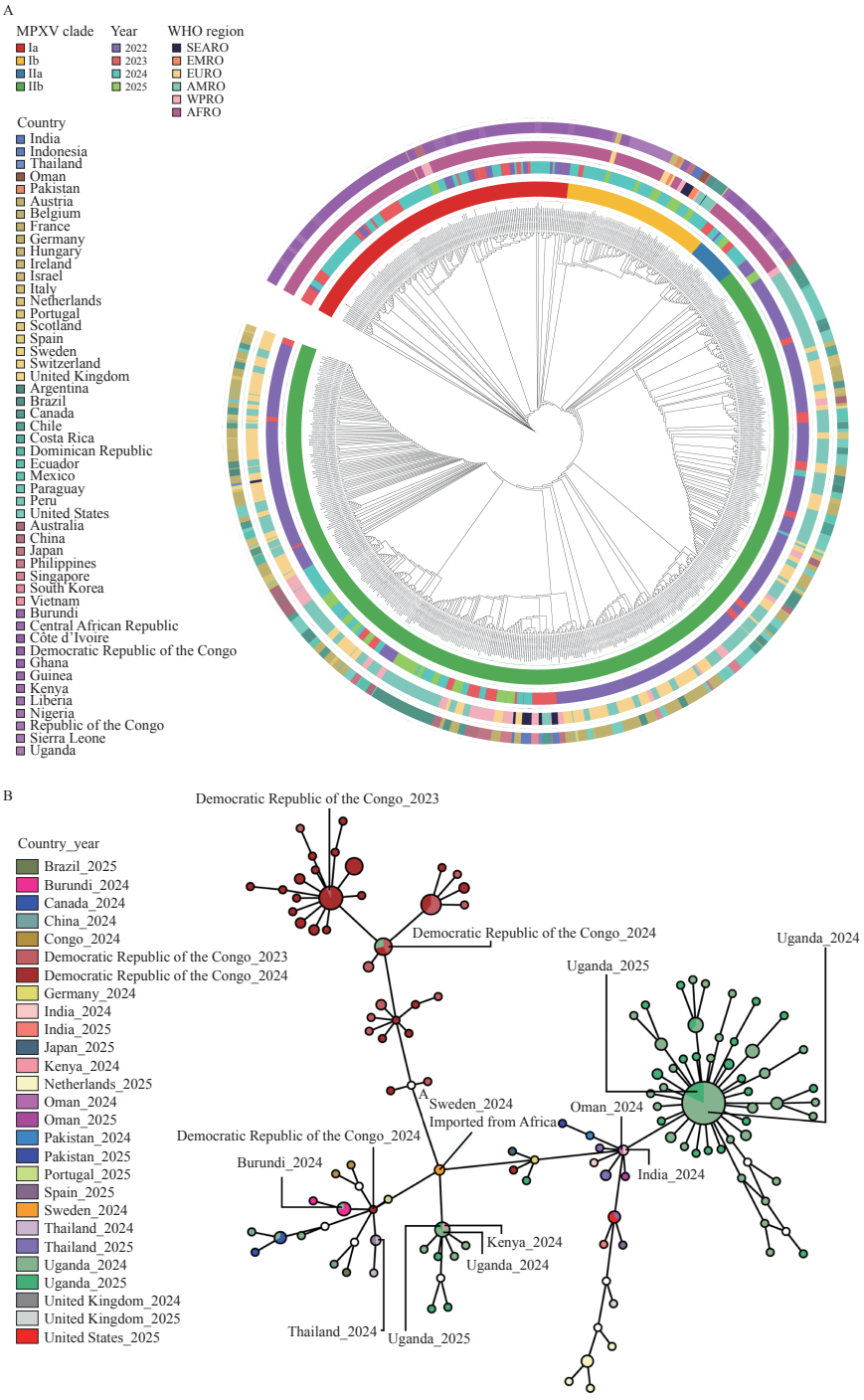


FIGURE 2. Phylogenetic analysis of four MPXV clades and haplotype network characterization of MPXV clade Ib. (A) Phylogenetic tree constructed from complete genomes of four MPXV clades; (B) Haplotype network analysis of MPXV clade Ib (January 2022 to September 2025).

Note: For (A), a total of 14,089 MPXV whole-genome sequences available from GISAID during the study period were retrieved. The sequences were filtered according to the following criteria: 1) genome length ≥190,000 base pairs; 2) absence of "warn/alert" flags; 3) no insertions or gaps ≥200 nucleotides; 4) absence of "stop/truncation" annotations; and 5) only countries with >50 cumulative mpox cases were included. After a multistep sampling procedure and removal of sequences with genome coverage <95% (Supplementary Table S2, available at https://weekly.chinacdc.cn/), the final dataset comprised 730 sequences, including 183 clade Ia, 45 clade Ib, 10 clade IIa, and 492 clade IIb genomes. For (B) MPXV clade Ib whole-genome sequences available from GISAID during the study period were retrieved. Sequences with genome coverage >95% were retained. A total of 245 sequences were included for haplotype network construction.

Abbreviation: MPXV=mpox virus.

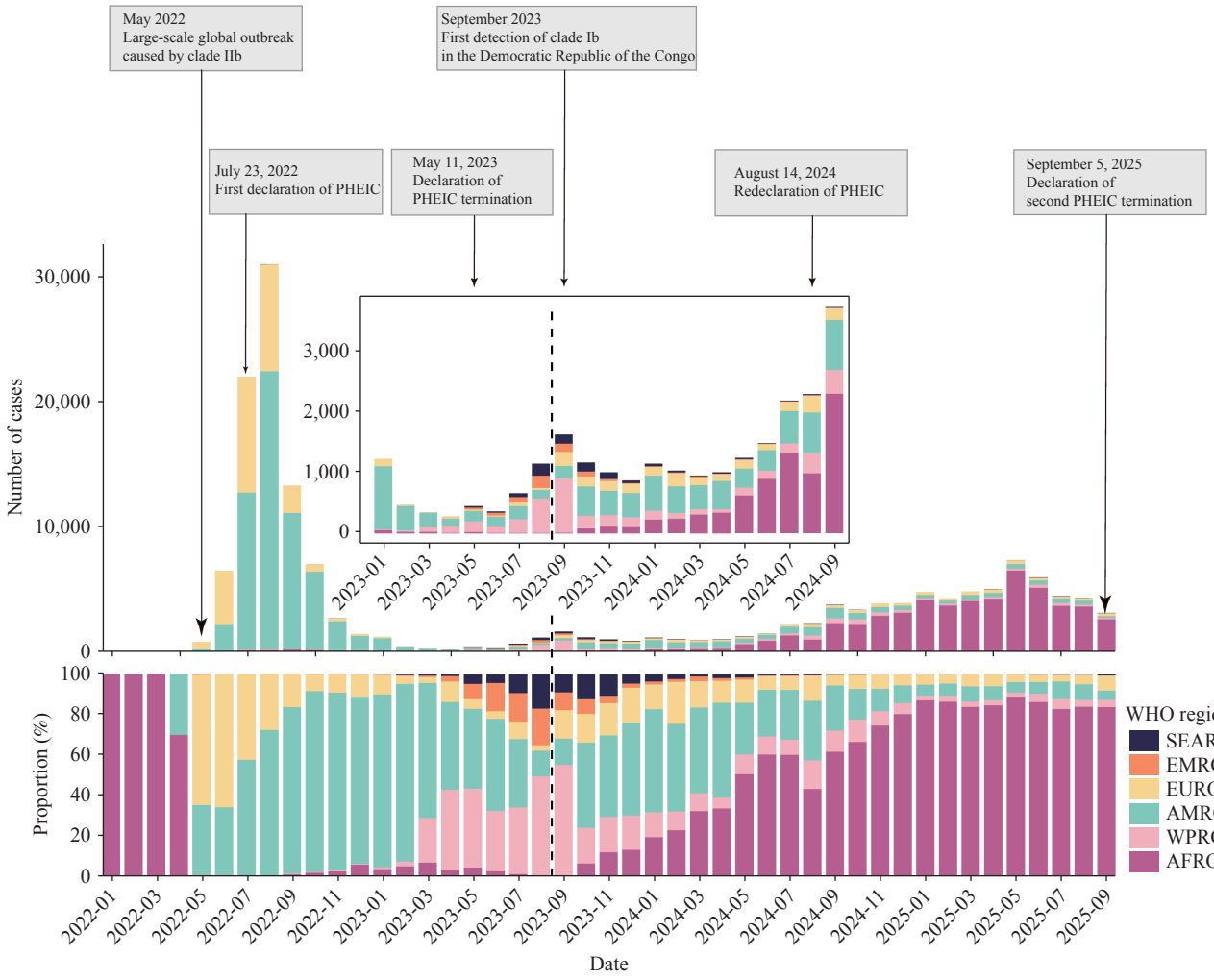
Innovation Center for Philosophy and Social Sciences, Beijing, China; [6] State Key Laboratory of Respiratory Health and Multimorbidity, Beijing, China; [7] Key Laboratory of Pathogen Infection Prevention and Control (Peking Union Medical College), Ministry of Education, Beijing, China.
[&] Joint first authors.

Submitted: January 19, 2026
Accepted: April 07, 2026
Issued: May 15, 2026

REFERENCES

1. Shan KJ, Wu CC, Tang XL, Lu RJ, Hu YL, Tan WJ, et al. Molecular evolution of protein sequences and codon usage in Monkeypox viruses. Genom Proteom Bioinf 2024;22(1):qzad003. https://doi.org/10.1093/gpbjnl/qzad003.
2. WHO. 2022-24 Mpox (Monkeypox) outbreak: Global Trends. 2024. https://worldhealthorg.shinyapps.io/mpx_global/. [2025-10-31].
3. Isidro J, Borges V, Pinto M, Sobral D, Santos JD, Nunes A, et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. Nat Med 2022;28(8):1569 – 72. https://doi.org/10.1038/s41591-022-01907-y.
4. McQuiston JH, Luce R, Kazadi DM, Bwangandu CN, Mbala-Kingebeni P, Anderson M, et al. U.S. preparedness and response to increasing clade I Mpox cases in the democratic republic of the Congo - United States, 2024. MMWR Morb Mortal Wkly Rep 2024;73(19):435-40. http://dx.doi.org/10.15585/mmwr.mm7319a3.
5. Brosius I, Vakaniaki EH, Mukari G, Munganga P, Tshomba JC, De Vos E, et al. Epidemiological and clinical features of mpox during the clade Ib outbreak in South Kivu, Democratic Republic of the Congo: a prospective cohort study. Lancet 2025;405(10478):547 – 59. https://doi.org/10.1016/S0140-6736(25)00047-9.
6. Vakaniaki EH, Kacita C, Kinganda-Lusamaki E, O'Toole A, Wawina-Bokalanga T, Mukadi-Bamuleka D, et al. Sustained human outbreak of a new MPXV clade I lineage in eastern Democratic Republic of the Congo. Nat Med 2024;30(10):2791 – 5. https://doi.org/10.1038/s41591-024-03130-3.
7. Otieno JR, Ruis C, Onoja AB, Kuppalli K, Hoxha A, Nitsche A, et al. Global genomic surveillance of monkeypox virus. Nat Med 2025;31(1):342 – 50. https://doi.org/10.1038/s41591-024-03370-3.
8. Borges V, Duque MP, Martins JV, Vasconcelos P, Ferreira R, Sobral D, et al. Viral genetic clustering and transmission dynamics of the 2022 mpox outbreak in Portugal. Nat Med 2023;29(10):2509 – 17. https://doi.org/10.1038/s41591-023-02542-x.
9. Evers YJ, Dukers-Muijrers NHTM, Van Liere GAFS, Hoebe CJPA. Sex abroad among men who have sex with men and its association with Chemsex, sexual risk behavior, and sexually transmitted diseases: a cross-sectional study in the Netherlands. Sex Transm Dis 2020;47(9):e29 – 32. https://doi.org/10.1097/OLQ.0000000000001207.
10. Qiao HJ, Paansri P, Escobar LE. Global Mpox spread due to increased air travel. Geospat Health 2024;19(1):10. https://doi.org/10.4081/gh.2024.1261.
11. Bangwen E, Diavita R, De Vos E, Vakaniaki EH, Nundu SS, Mutombo A, et al. Suspected and confirmed mpox cases in DR Congo: a retrospective analysis of national epidemiological and laboratory surveillance data, 2010-23. Lancet 2025;405(10476):408 – 19. https://doi.org/10.1016/S0140-6736(24)02669-2.
12. Ndembi N, Abdool Karim SS. Africa aims to avert an mpox pandemic. Science 2024;386(6717):7. https://doi.org/10.1126/science.adt4793.
13. Chakraborty C, Bhattacharya M, Sharma AR, Dhama K. Evolution, epidemiology, geographical distribution, and mutational landscape of newly emerging monkeypox virus. GeroScience 2022;44(6):2895 – 911. https://doi.org/10.1007/s11357-022-00659-4.
14. Sulaiman SK, Isma'il Tsiga-Ahmed F, Musa MS, Makama BT, Sulaiman AK, Abdulaziz TB. Global prevalence and correlates of mpox vaccine acceptance and uptake: a systematic review and meta-analysis. Commun Med 2024;4(1):136. https://doi.org/10.1038/s43856-024-00564-1.
15. Adetifa I, Muyembe JJ, Bausch DG, Heymann DL. Mpox neglect and the smallpox niche: a problem for Africa, a problem for the world. Lancet 2023;401(10390):1822 – 4. https://doi.org/10.1016/S0140-6736(23)00588-3.
16. Yu S, Wang HB, Wang DS, Dai SM, Sun S, Zhang JX, et al. Whole-genome sequencing and phylogenetic analysis of the first and second confirmed cases of human infection with Monkeypox virus from Sierra Leone. J Infect 2025;91(3):106587. https://doi.org/10.1016/j.jinf.2025.106587.

SUPPLEMENTARY MATERIAL



SUPPLEMENTARY FIGURE S1. Monthly trends and regional composition of newly reported mpox cases across the six WHO regions from January 2022 to September 2025.
Note: Six colors represent WHO regions: SEARO, EMRO, EURO, AMRO, WPRO, and AFRO.
Abbreviation: WHO=World Health Organization; SEARO=South-East Asia Region; EMRO=Eastern Mediterranean Region; EURO=European Region; AMRO=Region of the Americas; WPRO=Western Pacific Region; AFRO=African Region.

SUPPLEMENTARY TABLE S1. Comparison of the epidemiological characteristics of mpox between two epidemic phases, from January 2022 to August 2023, and from September 2023 to September 2025.

Epidemiological characteristic	The first phase (January 2022 to August 2023)	The second phase (September 2023 to September 2025)
Dominant pathogen	Clade IIb	Clade Ib
Onset of epidemic increase	May 2022	March 2024
Peak period	August 2022	May 2025
Peak case count	31,008 cases	7,343 cases
Cumulative cases	89,870 cases	75,374 cases
Predominant epidemic regions (proportion)	Region of the Americas and European Region (96.18%)	African Region (71.48%)

SUPPLEMENTARY TABLE S2. Sampling procedure and dataset of genomic sequences used for phylogenetic analysis.

MPXV clades	Total	Round 1 (Random sampling approximately 20%)	Round 2 (20% random sampling from clade IIb, others unchanged)	Round 3 (Removal of sequences with coverage less than 95%)
Clade Ia	1,049	210	210	183
Clade Ib	589	118	118	45
Clade IIa	39	10	10	10
Clade IIb	12,412	2,482	496	492
Total	14,089	2,820	834	730

SUPPLEMENTARY TABLE S3. Number of sequences and reported cases in countries included in the analysis.

Country	Number of sequences included in this study	Reported case count	Included sequences/ case count, (%)
Argentina	2	1,376	0.15
Australia	13	1,769	0.73
Austria	3	381	0.79
Belgium	3	907	0.33
Brazil	39	14,328	0.27
Burundi	14	4,483	0.31
Canada	32	2,236	1.43
Central African Republic	9	180	5.00
Chile	5	1,677	0.30
China	10	3,618	0.28
Costa Rica	2	232	0.86
Côte d'Ivoire	5	167	2.99
Democratic Republic of the Congo	203	30,615	0.66
Dominican Republic	3	83	3.61
Ecuador	4	629	0.64
France	3	4,539	0.07
Germany	95	4,532	2.10
Ghana	5	785	0.64
Guinea	2	1,040	0.19
Hungary	1	88	1.14
India	2	51	3.92
Indonesia	6	88	6.82
Ireland	17	320	5.31

Continued

Country	Number of sequences included in this study	Reported case count	Included sequences/ case count, (%)
Israel	1	378	0.26
Italy	12	1,233	0.97
Japan	11	256	4.30
Kenya	2	624	0.32
Liberia	2	940	0.21
Malawi	2	122	1.64
Mexico	10	4,862	0.21
Netherlands	8	1,557	0.51
Nigeria	4	1,384	0.29
Oman	2	7	28.57
Pakistan	2	29	6.90
Paraguay	3	130	2.31
Peru	24	3,959	0.61
Philippines	9	334	2.69
Portugal	13	1,255	1.04
Republic of the Congo	12	128	9.38
Sierra Leone	2	5,345	0.04
Singapore	2	81	2.47
Republic of Korea	6	188	3.19
Spain	5	9,028	0.06
Sweden	2	373	0.54
Switzerland	2	651	0.31
Thailand	7	960	0.73
Uganda	20	8,209	0.24
United Kingdom	10	4,354	0.23
USA	148	35,590	0.42
Vietnam	1	211	0.47
Total	800	156,312	0.51

Preplanned Studies

A Comprehensive Comparison of Transition Point Detection Methods for Monkeypox — Africa, 2024–2025

Mengshou Wang[1]; Zibin Yu[1]; Ling Yin[2,#]; Liu Hong[1,#]

Summary

What is already known about this topic?

Since May 2022, the global outbreak of mpox has presented a substantial public health challenge, with recent severe resurgences reported in several African countries, particularly in the Democratic Republic of the Congo (DRC).

What is added by this report?

This study developed compartmental- and network-based models using epidemiological data from the DRC, a country that has experienced a recent major outbreak. The applicability of the four methods for transition point detection in mpox transmission was confirmed for the first time. The strengths and limitations of each approach were evaluated systematically and quantitatively.

What are the implications for public health practice?

These findings provide theoretical guidance for identifying epidemic transition points in the mpox outbreak response and offer a reference for the rational allocation of public health resources during epidemics.

ABSTRACT

Introduction: Global mpox outbreaks pose a significant public health challenge. Accuracy predicting transition points can guide containment measures. However, different forecasting methods identified distinct types of transition points, making it essential to clarify their applicability.

Methods: Using World Health Organization (WHO) mpox outbreak data, this study employed both compartmental epidemic models and network models for simulation. Four transition point detection methods, dynamic network biomarkers (DNB), Network Information Gain (NIG), Temporal Change Point Detection (TCD), and Spatial-Temporal Principal Component Analysis (stPCA), were applied for predictions.

Results: The study revealed that NIG- and DNB-based methods were more suitable for detecting the peak timing of new cases, the epidemiological inflection point, while TCD and stPCA were more applicable for identifying the point of maximum acceleration in case incidence. These findings were validated in three African countries with high outbreak intensities too.

Conclusion: NIG and DNB can predict the peak timing of new mpox cases, whereas TCD and stPCA are suitable for forecasting the point of maximum acceleration in case incidence. These approaches enable more rational allocation of epidemic control resources and timely adjustment of containment measures, offering critical insights for managing mpox outbreaks.

Since May 2022, WHO-reported mpox outbreaks have become a global public health challenge (*1*), with severe resurgences in Africa threatening human health and well-being. Correctly identifying the epidemiological transition from a low to high prevalence deepens understanding of outbreak dynamics and enables more effective control strategies.

Utilizing data from the WHO "Global Mpox Trends" report and population-normalized confirmed case counts from Our World in Data (*2–3*), this study performed a comprehensive comparison of four state-of-the-art transition point detection methods for a representative country, the Democratic Republic of the Congo (DRC), from January 1, 2024, to November 30, 2025. These methods include NIG (*4*), TCD (*5*), stPCA (*6*), and DNB (*7*).

Given the mpox transmission patterns, it is reasonable to select the SEIR compartmental model, which incorporates an exposed (latent) period, for the analysis and discussion of the epidemic, as illustrated in Figure 1B. The SEIR model can be derived from a more detailed network model in which individuals serve as nodes and contacts serve as edges by imposing a homogeneous mixing assumption. Based on data from the WHO "Global Mpox Trends" report, we

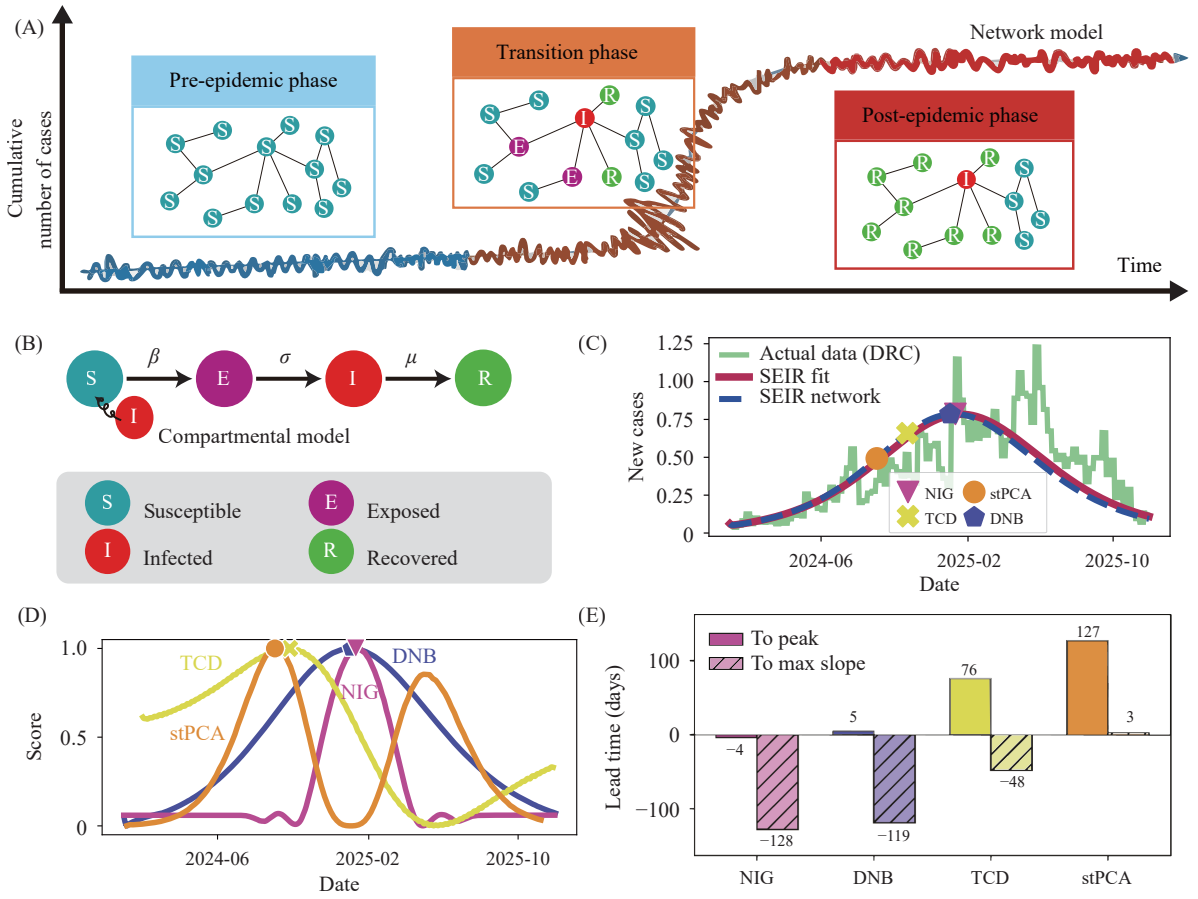


FIGURE 1. Epidemic transmission and phase transition of mpox in the DRC. (A) Schematic diagram of epidemic phase transition in a network model. (B) SEIR transmission model used for modeling. (C) Fitting of compartmental and network models to data from the DRC, with predictions from four transition-point detection methods. (D) Normalized scores and (E) time differences between the identified transition points in the network model and ground truth for the four methods, respectively.

Abbreviation: DRC=Democratic Republic of the Congo; SEIR=susceptible-exposed-infectious-recovered; NIG=network information gain; stPCA=spatial-temporal principal component analysis; TCD=temporal change-point detection; DNB=dynamic network biomarker.

derived the following epidemiological parameters for the DRC: transmission rate β=0.38, the reciprocal of the latent period σ=0.10, the reciprocal of the recovery period γ=0.33, and notably, the proportion of susceptible individuals in the total population, denoted as scale, was 1.2 ‰. Based on these parameters, a corresponding Watts-Strogatz (WS) network model was established with a mean degree of six and 2,000 nodes. The details are provided in the Supplementary Material. As shown in Figure 1C, both the compartmental and network models demonstrate consistent and good fits to the data.

We are now ready to apply these four methods to identify the critical transition point for the spread of mpox in the DRC, as depicted in Figure 1A. The predicted epidemic transition points derived from the four distinct methodologies are summarized as follows.

Both the NIG and DNB methods forecast the timing of the peak in daily newly reported case numbers, corresponding to the inflection point in cumulative cases. Specifically, the NIG prediction is closer to the actual peak than that of the DNB, although it lags by 4 days, while the DNB precedes the peak by 5 days. In contrast, the TCD and stPCA methods predicted a point much earlier than the peak of new daily cases. In fact, they estimated the time of maximum acceleration in case incidence, i.e. the point of the largest rate of increase in new daily cases, with deviations of 48 and 3 days, respectively. From this perspective, it can be concluded that the NIG and DNB methods primarily aimed to predict the timing of the peak in new daily cases, while the TCD and stPCA methods focused on predicting the point of maximum acceleration during case growth, as depicted in Figure 1E.

Figure 1D shows the normalized score plots for the four methods, where the time point corresponding to the peak position of each curve represents the predicted transition point. It can be observed that the score plots of NIG, DNB and TCD all exhibit a single distinct peak, except for the stPCA which shows two notable peaks with the second one corresponding to the inflection point during the decaying stage in daily new cases. Therefore, stPCA can capture additional transition signals that are not reflected by the other three methods.

These findings provide valuable guidelines for real-world applications. For the purpose of early warning of epidemics, the TCD and stPCA methods appear to be suitable choices, as they are more sensitive to the transition from a low level of prevalence to a high level, which is of critical importance for the design of effective epidemic prevention strategies. However, because the NIG and DNB methods are aligned with the peak of new cases or the inflection point of cumulative cases, they provide an important indicator of the strategy change in epidemic control from a strict policy to a relatively loose one.

To ensure the validity of our results under different scenarios, the aforementioned methods were applied to identify the mpox epidemic transition points in three other African countries: Uganda, Kenya, and Zambia. As shown in Figure 2, the results obtained are all consistent with those presented for the DRC, indicating that these four methods exhibit wide applicability and yield consistent outcomes across different regions.

DISCUSSION

Based on the above results, NIG and DNB belong to a category of methods predicting the timing of the peak in new case numbers, while TCD and stPCA fall into a category focused on predicting the time of maximum acceleration in case incidence. The main reason for this is the distinct logic of the four methods used for detecting the transition point. First, the NIG method, which focuses on network flow entropy and information gain, captures critical transitions by quantifying the modular interaction perturbations between an individual sample network and a reference

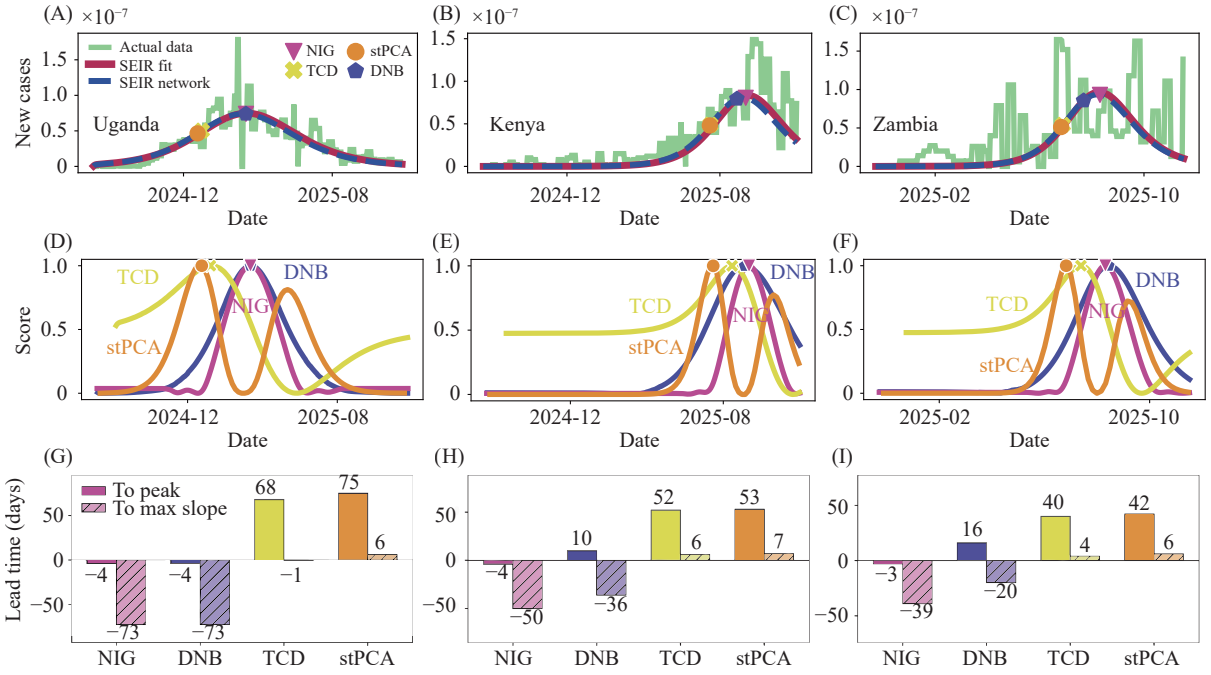


FIGURE 2. Results for three other severely affected countries in Africa: Uganda, Kenya, and Zambia. (A, B, C) Fitting of compartmental and network models to the data. Predicted transition points of four detection methods are marked by symbols. (D, E, F) Normalized scores predicted by four methods. The position of the largest score indicates the critical transition point. (G, H, I) Time difference between the identified transition points and the ground truth for the four methods respectively.

Abbreviation: DRC=Democratic Republic of the Congo; SEIR=susceptible-exposed-infectious-recovered; NIG=network information gain; stPCA=spatial-temporal principal component analysis; TCD=temporal change-point detection; DNB=dynamic network biomarker.

set. At the peak of new cases in the SEIR model within the WS network, the topological reconstruction of inter-node connections and changes in modular associations reached their most significant levels, with information differences at the network level accumulating to a maximum, thus enabling the NIG method to identify the tipping point. The DNB method relies on characteristics including collective fluctuations in local modules, enhanced internal correlations, and weakened external correlations in individual networks. When disease transmission reaches its peak in a WS network, the synergistic fluctuations of key node modules satisfy the core criteria for DNB identification, leading to synchronous detection of the critical state at the peak.

In contrast, the stPCA method, based on Takens' delay embedding theory, converts the high-dimensional spatial information of the WS network into univariate temporal dynamics, with its core focus on capturing the early evolutionary characteristics of the system's center manifold without waiting for significant mutations in the network topology. When the SEIR model propagated to the inflection point, the system exhibited dynamic trends prior to the peak, and early warning signals were captured in advance through spatiotemporal information conversion. Finally, the TCD method generates prediction sequences via randomly distributed embedding and quantifies subtle fluctuations in the prediction errors by integrating Bayesian online change-point detection. In the early stages of disease transmission in the WS network, subtle structural changes in node connections have already manifested through prediction errors, eliminating the need to wait for significant topological perturbations at the peak. Thus, the inflection point is detected when the number of new cases reaches half of the peak.

Our study is the first to confirm the applicability of these four methods for transition-point detection in mpox transmission. This conclusion does not rely on the specific details of the SEIR or WS network models, nor is it constrained by the reported mpox data from the DRC, Uganda, Kenya, and Zambia. For example, the DNB method has been used to detect the seasonal influenza outbreaks in Tokyo, Japan (*8*).

The generalizability of our findings in practice is subject to the following limitations:

First, the effectiveness of the methods for other infectious diseases requires further validation, the homogeneous mixing assumption may not accurately represent structured contact patterns, and potential underreporting in some African countries may affect data comprehensiveness.

Second, the type and scale of the network, variations in data sources, and deviations between the actual acquired data and ground truth owing to noise interference could all exert a non-negligible impact on the actual performance of the method in real-world applications.

Third, due to data availability, the current study relies solely on data from African countries, which may introduce geographical bias.

Fourth, only recent viral strains were considered, which leads to potential limitations in strain representativeness. We also acknowledge the limitations of using lead time as a single evaluation metric, which may not sufficiently distinguish between the different implications of early and delayed predictions in real epidemic prevention and control practices.

In summary, through a discussion of four transition-point detection methods (NIG, TCD, stPCA, and DNB) in the context of the mpox outbreak, this study clarified their respective applicable scopes. These approaches facilitate a more rational allocation of epidemic control resources and the timely adjustment of containment measures, thereby providing important insights for the management of mpox outbreaks.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the National Key R&D Program of China (Grant No. 2023YFC2308702).

doi: 10.46234/ccdcw2026.099

[#] Corresponding authors: Ling Yin, yinling@siat.ac.cn; Liu Hong, hongliu@sysu.edu.cn.

[1] School of Mathematics, Sun Yat-sen University, Guangzhou City, Guangdong Province, China; [2] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen City, Guangdong Province, China.

Submitted: January 20, 2026
Accepted: May 05, 2026
Issued: May 15, 2026

REFERENCES

1. Karagoz A, Tombuloglu H, Alsaeed M, Tombuloglu G, AlRubaish AA, Mahmoud A, et al. Monkeypox (mpox) virus: classification, origin, transmission, genome organization, antiviral drugs, and molecular diagnosis. J Infect Public Health 2023;16(4):531 – 41. https://doi.org/10.1016/j.jiph.2023.02.003.

2. World Health Organization. Global Mpox Trends. 2025. https://worldhealthorg.shinyapps.io/mpx_global/. [2025-12-19].
3. Mathieu E, Spooner F, Dattani S, Ritchie H, Roser X. Monkeypox. Our World in Data. 2022. https://ourworldindata.org/mpox. [2025-12-20].
4. Yan JL, Li PL, Li Y, Gao R, Bi C, Chen LN. Disease prediction by network information gain on a single sample basis. Fundam Res 2025;5(1):215 – 27. https://doi.org/10.1016/j.fmre.2023.01.009.
5. Hou JW, Ma HF, He DK, Sun J, Nie Q, Lin W. Harvesting random embedding for high-frequency change-point detection in temporal complex systems. Natl Sci Rev 2022;9(4):nwab228. https://doi.org/10.1093/nsr/nwab228.
6. Chen P, Suo YF, Aihara K, Li Y, Wu D, Liu R, et al. Ultralow-dimensionality reduction for identifying critical transitions by spatial-temporal PCA. Adv Sci 2025;12(20):2408173. https://doi.org/10.1002/advs.202408173.
7. Liu XP, Chang X, Leng SY, Tang H, Aihara K, Chen LN. Detection for disease tipping points by landscape dynamic network biomarkers. Natl Sci Rev 2019;6(4):775 – 85. https://doi.org/10.1093/nsr/nwy162.
8. Chen P, Chen E, Chen LN, Zhou XJ, Liu R. Detecting early-warning signals of influenza outbreak based on dynamic network marker. J Cell Mol Med 2019;23(1):395 – 404. https://doi.org/10.1111/jcmm.13943.

SUPPLEMENTARY MATERIALS

Network Model

Modeling infectious diseases at the individual level using complex network frameworks enables a more granular analysis of transmission dynamics and control measures. In this model, the nodes represent individuals, and the edges denote the contacts between them. Based on the mpox transmission mechanism, node states are categorized into four classes: susceptible individuals who have not yet been infected (S), exposed individuals who are infected but asymptomatic and non-infectious (E), infectious individuals who are symptomatic and capable of transmitting the disease (I), and recovered individuals who are immune to reinfection (R), as shown in Figure 1B.

Consider a network of N nodes indexed by $i=1,\ldots,N$, with adjacency matrix A. If node i is in contact with node j, then $a_{ij} = 1$; otherwise, $a_{ij} = 0$. Each node i can be in one of the states: S, E, I, or R. The transition probabilities per unit time are defined as follows: 1) A node in state S can be infected by each of its infectious neighbors (state I) with probability β, transitioning to state E; 2) A node in state E transitions to state I with probability σ; 3) A node in state I transitions to state R with probability γ.

Let $p_i^S(t), p_i^E(t), p_i^I(t)$, and $p_i^R(t)$ denote the probabilities that node i is in states S, E, I, and R at time t, respectively. Using a Markov state transition tree to derive the state probabilities at time $t + \delta t$ and taking the limit $\delta t \to 0$, we obtain the following system of master equations:

$$\begin{aligned}
\frac{dp_i^S(t)}{dt} &= -\beta p_i^S(t)\sum_j^N a_{ji}p_j^I(t)\\
\frac{dp_i^E(t)}{dt} &= \beta p_i^S(t)\sum_j^N a_{ji}p_j^I(t) - \sigma p_i^E(t)\\
\frac{dp_i^I(t)}{dt} &= \sigma p_i^E(t) - \gamma p_i^I(t)\\
\frac{dp_i^R(t)}{dt} &= \gamma p_i^I(t)
\end{aligned}$$

Further details can be found in reference (*1*). Based on the next-generation matrix method, the basic reproduction number can be derived as $R_0 = \rho(A)\beta/\gamma$, where $\rho(A)$ denotes the spectral radius, i.e. the largest eigenvalue, of matrix A.

Network models play a significant role in epidemiology, accurately forecasting epidemic transmission dynamics, analyzing viral spreading mechanisms, identifying super-spreaders and transmission hubs, guiding public health intervention strategies, optimizing non-pharmaceutical interventions, and designing targeted vaccination strategies.

Compartmental Model

For the network model, we adopted a homogeneous mean-field approximation by assuming that all the nodes were statistically identical. Define $S(t)=\sum_i p_i^S(t)/N$, $E(t) = \sum_i p_i^E(t)/N$, $I(t)=\sum_i p_i^I(t)/N$, and $R(t)=\sum_i p_i^R(t)/N$. Furthermore, the original adjacency matrix is replaced by a fully connected network with uniform link strength $a_{ij} = \frac{\langle k^2\rangle}{\langle k\rangle}$ or $a_{ij} = \langle k\rangle$, where $\langle k\rangle$ denotes the average degree of the network, but preserving the total number of edges. In the limit $N \to \infty$, the mean-field equations become:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\overline{\beta}S(t)I(t)\\
\frac{dE(t)}{dt} &= \overline{\beta}S(t)I(t) - \sigma E(t)\\
\frac{dI(t)}{dt} &= \sigma E(t) - \gamma I(t)\\
\frac{dR(t)}{dt} &= \gamma I(t)
\end{aligned}$$

where $\overline{\beta} = \frac{\langle k^2\rangle}{\langle k\rangle}\beta$ or $\overline{\beta} = \langle k\rangle\beta$, and the conservation law $S(t)+E(t)+I(t)+R(t)=1$ holds. This formulation is identical to the classical compartmental model. The basic reproduction number for this model is given by $R_0 = \overline{\beta}/\gamma$.

The compartment model has a long history in the study of epidemics dating back to the pioneering work of Kermack and McKendrick in 1927. It offers high interpretability and features a relatively simple, flexible, and expandable structure. The compartment model does not require complex or high-quality data, making it a widely applied mainstream model for forecasting infectious diseases. However, it highly depends on the researcher's understanding of the disease transmission mechanism, and the omission of network structure details greatly restricts its application.

In this study, we integrate both modeling approaches: the compartmental (SEIR) model provides foundational population-level estimates of transmission dynamics, while the network model simulates the heterogeneous, contact-driven spread within a structured population. This combination allows us not only to calibrate core epidemiological parameters, but also to identify early warning signals of accelerated transmission that precede detectable shifts in aggregate case counts.

Data and Parameters

Drawing on epidemiological parameter data from the World Health Organization (WHO) "Global Mpox Trends" released on December 18, 2025, and confirmed case counts from Our World in Data, we adopt the following assumptions: the incubation period follows an exponential distribution, giving $\sigma = 1/\mathbb{E}[\mathrm{L}]$, where $\mathbb{E}[\mathrm{L}]$ is the mean incubation period; transmission occurs after symptom onset and is uniformly random during the infectious period, leading to the approximation

$$\mathbb{E}(\mathrm{SI}) \approx \mathbb{E}(\mathrm{GT}) \approx \text{mean incubation period} + \frac{1}{2} \times \text{mean infectious period}$$

where $\mathbb{E}(\mathrm{SI})$ is the mean serial interval and $\mathbb{E}(\mathrm{GT})$ is the mean generation time. Consequently, $\gamma = 1/(2(\mathbb{E}(\mathrm{SI}) - \mathbb{E}[\mathrm{L}]))$.

The Democratic Republic of the Congo (DRC) was selected as the representative country, with data spanning from January 1, 2024, to November 30, 2025. The corresponding epidemiological parameters are $\mathbb{E}[\mathrm{L}] = 9.9$ days and $\mathbb{E}[\mathrm{SI}] = 11.4$ days (*2*), resulting in $\sigma = 0.10$ and $\gamma = 0.33$. Fitting the SEIR model to this dataset produces an estimated mean transmission rate $\bar{\beta} = 0.38$. A parameter scale was introduced during the model fitting process, which represents the proportion of susceptible individuals within the total population.

Based on the available data from global repositories (e.g., INRB, GISAID, and GenBank) and alignment with WHO reports on global mpox trends, the circulating virus strains in the DRC during the study period predominantly belonged to Clade I, with subclades Ia and Ib co-circulating across provinces. Because both subclades fell within the same major clade (Clade I), our analysis did not further differentiate their potential effects.

Methods for Transition Point Detection

Identifying the critical moments at which a system is about to undergo a drastic transition before a large-scale outbreak occurs is a core question in the field of epidemiology. Its value lies in providing key temporal references for assessing epidemic trends and informing public health decision making. Here, we list four typical methods used for transition-point detection in the literature.

Dynamic network biomarker (DNB): When a physical system approaches a critical state from a stable normal state, a DNB module (a group of nodes) appears and satisfies the following three requirements: 1) Deviation for each node inside the module drastically increases; 2) Correlation between nodes inside the module rapidly increases; 3) Correlation between nodes inside and outside of the module rapidly decreases (*3*). The DNB module is obtained by maximizing the DNB score.

$$\mathrm{I_{DNB}} = \mathrm{SD_{in}} \frac{\mathrm{PCC_{in}}}{\mathrm{PCC_{out}}},$$

where $\mathrm{SD_{in}}$ is the standard deviation of the DNB module, $\mathrm{PCC_{in}}$ is the average of the absolute Pearson correlation coefficients among the nodes in the DNB module, and $\mathrm{PCC_{out}}$ is the average of the nodes within and outside the DNB module.

Network information gain (NIG): The NIG records the global disturbed information gain between the reference samples and the one case sample (*4*). Therefore, it characterizes the network fluctuations of the nodes in a system rather than random fluctuations, which is the core of quantifying the criticality of the network. For a network of m

nodes, the NIG score is defined as the average difference between the cumulative information gains of n and n+1 samples respectively.

$$NIG = \frac{1}{m}\sum_{i=1}^{m}[IG_{n+1}(x_i) - IG_n(x_i)].$$

The cumulative information gain here is given by the difference between the node entropy $H_n(x_{ik})$ with x_{ik} as the central node and the conditional entropy $H_n(x_{ik}|x_i)$ when x_{ik} is considered a neighboring node of x_i:

$$IG_n(x_i) = \frac{1}{n_i}\sum_{k=1}^{n_i}[H_n(x_{ik}) - H_n(x_{ik}|x_i)],$$

where n_i denotes the total number of neighbors of node x_i. When the NIG score sharply increased and peaked, the system reached a transition point.

Spatial-temporal principal component analysis (stPCA): The stPCA is designed to maintain the interpretability of a linear subspace while representing a high-dimensional dynamical system with a single projected variable using a delay embedding strategy (*5*). For an observed high-dimensional time series X^t, one can construct a corresponding delayed vector Z^t by using a delay-embedding strategy, with $L > 1$ as the embedding dimension. They are related through the spatial-temporal information transformation (STI) equations $Z^t = \phi(X^t)$, whose linearization gives $Z^t = WX^t$. Similar to the PCA method, the loss function of the stPCA is given by

$$\min_w -(1-\lambda)\sum_{i=1}^{L}||W_iX||_2^2 + \lambda\sum_{i=1}^{L-1}||W_iP - W_{i+1}Q||_2^2,$$

where $X=[X^1,\ldots,X^m]$, $P=[X^2,\ldots,X^m]$, $Q=[X^1,\ldots,X^{m-1}]$, and $\lambda \in [0,1]$ is a regularization weight.

To identify the transition point of the dynamic system, a sliding window scheme was implemented that divided the entire process into smaller windows. Then, for each window, the high-dimensional time series were reduced to a one-dimensional trajectory of the latent variable using stPCA. Finally, the abrupt increase in the standard deviation provides a clear signal of the upcoming critical transition.

Temporal change-point detection (TCD): This approach first uses a Randomly Distributed Embedding (RDE) framework to compute predictions based on the time points in a prescribed sliding window of short length (*6*). Therefore, when the underlying system undergoes an abrupt state transition, the fidelity of the prediction is expected to decrease quickly, but then recover shortly after the small window slides through the change point. With respect to the standard deviation of the predictions and the prediction error, the BOCD test was applied, which aims to estimate the run-length posterior distribution and the posterior predictive distribution with Bayesian inference. Based on the run length distribution, a change point can be detected when the run length decreases to zero.

Computational Workflow

The computational workflow for this epidemiological modeling study is structured into four consecutive stages:

Data acquisition and preprocessing: The smoothed time series data for daily new cases per million in the Democratic Republic of the Congo were sourced from the Our World in Data database, serving as the empirical proxy for the SEIR model term $(\sigma \cdot E) \times \text{scale}$.

Parameter estimation via model calibration: A deterministic SEIR compartmental model was calibrated to this data by minimizing the Mean Squared Error (MSE) loss, thereby estimating the key epidemiological parameters (β, σ, γ), initial conditions (S_0, E_0, I_0, R_0), and the data scaling factor (scale).

Network model construction: Utilizing the parameter set derived from the preceding calibration, a corresponding network-based epidemic model was constructed to incorporate the population structure and contact heterogeneity.

Critical Transition Analysis: Finally, the transition point (or epidemic threshold) within the constructed network model was detected and compared using four distinct computational methods: Dynamic Network Biomarkers, Network Information Gain, Temporal Change-Point Detection, and Spatial-Temporal Principal Component Analysis.

Code is available at https://github.com/alierzy/Mpox.git.

REFERENCES

1. Wang MS, Peng LR, Jia BG, Hong L. Epidemic dynamics in physical-information-social multilayer networks. Phys A Stat Mech Appl 2025;678:130944. https://doi.org/10.1016/j.physa.2025.130944.
2. Marziano V, Guzzetta G, Longini I, Merler S. Epidemiologic quantities for monkeypox virus clade I from historical data with implications for current outbreaks, Democratic Republic of the Congo. Emerg Infect Dis 2024;30(10):2042 – 6. https://doi.org/10.3201/eid3010.240665.
3. Liu XP, Chang X, Leng SY, Tang H, Aihara K, Chen LN. Detection for disease tipping points by landscape dynamic network biomarkers. Natl Sci Rev 2019;6(4):775 – 85. https://doi.org/10.1093/nsr/nwy162.
4. Yan JL, Li PL, Li Y, Gao R, Bi C, Chen LN. Disease prediction by network information gain on a single sample basis. Fundam Res 2025;5(1):215 – 27. https://doi.org/10.1016/j.fmre.2023.01.009.
5. Chen P, Suo YF, Aihara K, Li Y, Wu D, Liu R, et al. Ultralow-dimensionality reduction for identifying critical transitions by spatial-temporal PCA. Adv Sci 2025;12(20):2408173. https://doi.org/10.1002/advs.202408173.
6. Hou JW, Ma HF, He DK, Sun J, Nie Q, Lin W. Harvesting random embedding for high-frequency change-point detection in temporal complex systems. Natl Sci Rev 2022;9(4):nwab228. https://doi.org/10.1093/nsr/nwab228.

Methods and Applications

Assessment of Cross-Border Monkeypox Importation Risk: A Modeling Study — Worldwide, 2022–2025

Qiangru Huang[1,&]; Xiaoli Wang[2,3,&]; Xuefeng Huang[4,&]; Dazhu Huo[1]; Mingyue Pan[2]; Ting Zhang[1,4]; Xinli Jiang[1]; Hualei Xin[1]; Zhongjie Li[1,4,5,6,#]; Weizhong Yang[1,4,5,6]

ABSTRACT

Introduction: Monkeypox (mpox) has re-emerged as a significant global public health threat. This study developed a practical, rapid early-warning tool to quantify country-level mpox importation risk and identify potential sources of cross-border transmission.

Methods: We obtained weekly confirmed mpox case data and international passenger volume data spanning January 2022 through November 2025. We constructed an importation risk index (RI) that integrated epidemic intensity in source countries, passenger flows, and public health response capacity. Using a simplified ordinary differential equation framework, we derived a regional RI to estimate time-varying growth and acceleration parameters from surveillance time series. We then computed both real-time and baseline RI values and evaluated their lagged correlations with reported case counts.

Results: The RI demonstrated consistently strong lagged correlations with subsequent reported case counts. The real-time RI provided earlier alerts in 93.3% of countries, with a mean lead time of 4.23 weeks — approximately 3.7 weeks ahead of the baseline RI. The dominant sources of importation risk varied across countries, suggesting that this model can inform tailored, country-specific prevention and control strategies.

Conclusion: This cross-border risk assessment model provides an actionable framework for early warning and source attribution, supporting targeted surveillance and proactive allocation of public health resources. These findings underscore the critical importance of global health collaboration and timely data sharing.

Monkeypox (mpox) is caused by the monkeypox virus (MPXV) (*1*). Sustained human-to-human transmission fueled a multinational outbreak, prompting the World Health Organization (WHO) to declare a Public Health Emergency of International Concern (PHEIC) in both 2022 and 2024 (*2*). The ongoing evolution of MPXV poses a persistent epidemic risk, necessitating continuous monitoring of viral mutations alongside the development of practical risk-warning tools (*3*–*4*).

Previous studies have assessed cross-border transmission risk using case data and passenger volumes between regions (*5*–*7*). However, delays in clinical consultation, sample processing, and case reporting frequently hinder timely public health responses. To address these gaps, we constructed a risk assessment model integrating airline network connectivity and healthcare control capacity to identify high-risk regions for mpox importation and enhance the speed and accuracy of risk assessments.

METHODS

Data Source and Preprocessing

Weekly confirmed mpox case data from January 2022 to November 2025 were obtained from the WHO dashboard (*8*). Monthly international air passenger data covering the same period were sourced from the Official Aviation Guide (OAG, www.oag.com).

Importation risk primarily originates from regions with high levels of epidemic activity. Countries with cumulative confirmed case counts exceeding 200 during the study period were selected for analysis; 48 countries met this criterion.

Risk Index Assessment

To address lags in data availability, we analyzed the importation risk index (RI) and its relationship to reported case counts. The RI is defined as follows:

$$RI_t : D_t \times F_t \times S_t \quad (1)$$

where D_t represents the weekly confirmed mpox case count reported by the source country at week t, normalized using min-max normalization across all source countries for each week to ensure cross-country comparability; F_t denotes the outbound air passenger volume from the source country to the destination country; S_t represents the stringency index of prevention and control measures, derived from the Global Health Security (GHS) Index — the GHS score was inverted (100–GHS) so that higher values indicate weaker health security capacity (*9*); and t is the week of assessment. $RI(d, f, s)$ thus quantifies the risk of cross-border mpox importation.

For temporal harmonization, weekly confirmed case data served as the base temporal resolution. Monthly international air passenger data were uniformly disaggregated to the weekly level. The stringency index, a relatively stable country-level indicator, was assigned as each country's most recent available score. Where case or passenger data were unavailable due to reporting delays, linear interpolation was applied to fill gaps.

Given the observed outbreak time-series data $d \in D$ for a region, we defined the regional risk index (RRI) at time t as $RRI = \cdot v(t)$, where $v(t) \in R^n$ is an n-dimensional vector representing multidimensional epidemic features of the region at time t, and $\in R^n$ is a weight vector balancing the contribution of each dimension to the RI. Equal weights were assigned to all dimensions. This parsimonious approach avoids subjective bias in the absence of empirical evidence favoring any single component and, because these dimensions capture complementary aspects of epidemic risk, provides a transparent and readily reproducible baseline.

To compute $v(t)$ explicitly, we constructed a simplified model guided by the principles of dynamic transmission modeling. Let $N(t)$ denote the number of infectious cases at time t in the source countries. We modeled $N(t)$ using the following ordinary differential equation:

$$\frac{dN(t)}{dt} = \lambda(t)N(t) \tag{2}$$

Here, $\lambda(t)$ is a time-varying coefficient encompassing the growth rate, control intensity, and effective reproduction number. A first-order linear approximation of $\lambda(t)$ was adopted for computational tractability and interpretability:

$$\lambda(t) = at + b \tag{3}$$

Within a sliding window of several weeks, the time-varying transmission coefficient can be reasonably approximated by a linear trend, as short-term changes in transmission intensity are generally monotonic. Parameter a captures the acceleration (or deceleration) of transmission, while b captures the baseline growth rate at the start of the window. Both parameters a and b were estimated from the time-series data ($d \in D$). Setting $\lambda(t^*) = 0$ yields the epidemic inflection point $t^* = -b/a$. Parameter a describes the acceleration of the infection rate under the combined influence of various factors. Denoting the total population of the region as P, we defined the following four-dimensional vector to characterize the epidemic risk situation:

$$v(t) = \left(\left| t - t^* \right|, \lambda(t), N(t), \frac{N(t)}{P} \right) \tag{4}$$

Lagged Correlation Assessment

The real-time RI is defined as the index computed using current case counts and passenger volume data, representing the best estimate under ideal data availability. The baseline RI (base RI) is defined as the index calculated under realistic reporting-delay conditions, with case data lagged by 4 weeks and air passenger data by 8 weeks, reflecting the data availability typical of routine surveillance settings.

We assessed the lagged correlation between the RI and reported case counts. Lead time was determined using an Exponentially Weighted Moving Average (EWMA) early-warning algorithm and was defined as the time difference (in weeks) between the RI triggering an alert and the reported case count triggering an alert.

RESULTS

We evaluated model effectiveness across three dimensions: risk index values, lag time, and primary sources of risk. Figure 1 shows that the mpox importation RI was highest in 2022, with peak values exceeding 1.0 in seven countries, primarily concentrated in Europe. During 2023 and 2024, peak RI values exceeded 1.0 in four and five countries, respectively, and the risk shifted toward East Asia. By 2025, overall importation risk remained low.

Among the 30 countries with the highest importation RI, reported case counts exhibited strong lagged correlations with RI values: 100% and 70.3% of countries showed correlation coefficients above 0.7 for

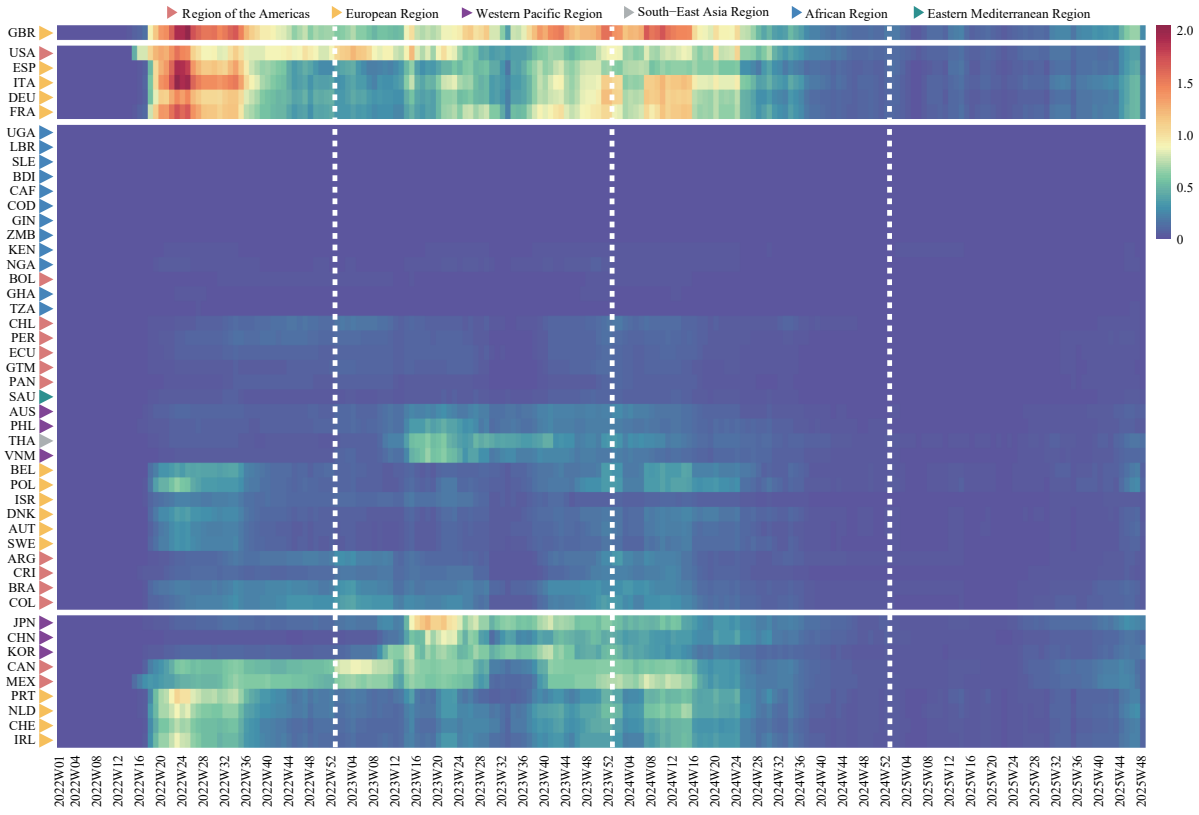


FIGURE 1. RI heatmap, 2022W01–2025W48.
Note: Forty-eight countries with cumulative cases greater than 200 are shown. The colors behind the row names correspond to different regions. White dashed lines separate different years.
Abbreviation: RI=risk index; ARG=Argentina; AUS=Australia; AUT=Austria; BDI=Burundi; BEL=Belgium; BOL=Bolivia; BRA=Brazil; CAF=Central African Republic; CAN=Canada; CHE=Switzerland; CHL=Chile; CHN=China; COD=Democratic Republic of the Congo; COL=Colombia; CRI=Costa Rica; DEU=Germany; DNK=Denmark; ECU=Ecuador; ESP=Spain; FRA=France; GBR=United Kingdom; GHA=Ghana; GIN=Guinea; GTM=Guatemala; IRL=Ireland; ISR=Israel; ITA=Italy; JPN=Japan; KEN=Kenya; KOR=Republic of Korea; LBR=Liberia; MEX=Mexico; NGA=Nigeria; NLD=Netherlands; PAN=Panama; PER=Peru; PHL=Philippines; POL=Poland; PRT=Portugal; SAU=Saudi Arabia; SLE=Sierra Leone; SWE=Sweden; THA=Thailand; TZA=United Republic of Tanzania; UGA=Uganda; USA=United States; VNM=Vietnam; ZMB=Zambia.

the real-time and baseline RIs, respectively (Table 1).

The real-time RI identified importation risks before a rapid increase in case counts in 93.3% of countries, with an average lead time of 4.23 weeks; by comparison, the baseline RI achieved a lead time of only 0.5 weeks (Figure 2). These findings demonstrate that both the real-time and baseline RIs can serve as effective early-warning tools, while underscoring the critical importance of timely surveillance data.

For most countries, importation risk was primarily driven by the United States, Australia, Spain, the Democratic Republic of the Congo, the United Kingdom, and Brazil (Figure 3). The dominant sources varied across recipient countries, indicating that this model can inform tailored, country-specific prevention and control strategies.

DISCUSSION

This risk assessment model offers a practical approach to quantifying and anticipating cross-border mpox importation risk by integrating epidemic intensity in source countries, cross-border passenger volumes, and contextual indicators of response capacity. The strong concordance between estimated importation risk and subsequent case trajectories supports the model's validity and robustness. This consistency was evident in both the trends and absolute levels of RI changes, confirming the reliability and stability of the risk assessment framework (*10*).

The cross-country variation in real-time RI lead time (0–15 weeks) reflects several interacting factors. Countries with well-established surveillance systems

TABLE 1. Lagged correlation between the importation risk index and reported cases, and the lead time for early detection of the increase in mpox cases.

No.	Country	Real-time risk index		Baseline risk index	
		r	Lead time (week)	*r*	Lead time (week)
1	ARG	0.96	9	0.96	4
2	AUT	0.99	3	0.86	0
3	BEL	0.99	2	0.69	0
4	BRA	0.99	3	0.84	0
5	CAN	0.95	0	0.31	−4
6	CHE	0.98	2	0.74	0
7	CHL	0.98	5	0.97	0
8	CHN	0.88	3	0.76	0
9	COL	0.99	7	0.99	1
10	CRI	0.83	15	0.41	12
11	DEU	0.98	0	0.46	−2
12	DNK	0.95	3	0.71	0
13	ECU	0.95	11	0.94	5
14	ESP	0.98	2	0.71	0
15	FRA	0.99	3	0.81	0
16	GBR	0.98	1	0.50	−3
17	GTM	0.89	14	0.89	10
18	IRL	0.90	4	0.87	0
19	ISR	0.95	2	0.63	0
20	ITA	0.96	1	0.72	0
21	JPN	0.86	4	0.84	0
22	KOR	0.87	8	0.89	0
23	MEX	0.98	7	0.97	1
24	NLD	0.98	1	0.52	−2
25	PER	0.99	3	0.92	−1
26	POL	0.98	5	0.97	−1
27	PRT	0.91	1	0.49	−1
28	SWE	0.90	2	0.74	0
29	THA	0.79	3	0.68	−2
30	USA	0.98	3	0.86	−2

Note: *r*, the optimal Pearson correlation coefficient from the lagged cross-correlation between the importation risk indices (real-time and baseline) and weekly reported case counts; lead time (weeks), the time difference between the RI triggering an early-warning alert and the reported case count triggering an alert, determined by the EWMA algorithm; positive values indicate the RI provided an earlier alert, negative values indicate a delayed alert.

Abbreviation: RI=risk index; EWMA=exponentially weighted moving average; mpox=monkeypox; ARG=Argentina; AUT=Austria; BEL=Belgium; BRA=Brazil; CAN=Canada; CHE=Switzerland; CHL=Chile; CHN=China; COL=Colombia; CRI=Costa Rica; DEU=Germany; DNK=Denmark; ECU=Ecuador; ESP=Spain; FRA=France; GBR=United Kingdom; GTM=Guatemala; IRL=Ireland; ISR=Israel; ITA=Italy; JPN=Japan; KOR=Republic of Korea; MEX=Mexico; NLD=Netherlands; PER=Peru; POL=Poland; PRT=Portugal; SWE=Sweden; THA=Thailand; USA=United States.

and short reporting delays (e.g., Germany and the United Kingdom) exhibited shorter lead times because their baseline RI already approximates real-time conditions, limiting the incremental predictive gain. Additionally, countries with diverse, multi-source travel connectivity may aggregate risk signals across multiple importation pathways, enabling earlier detection (*11*). These findings underscore the importance of reducing surveillance reporting delays to maximize early-warning performance.

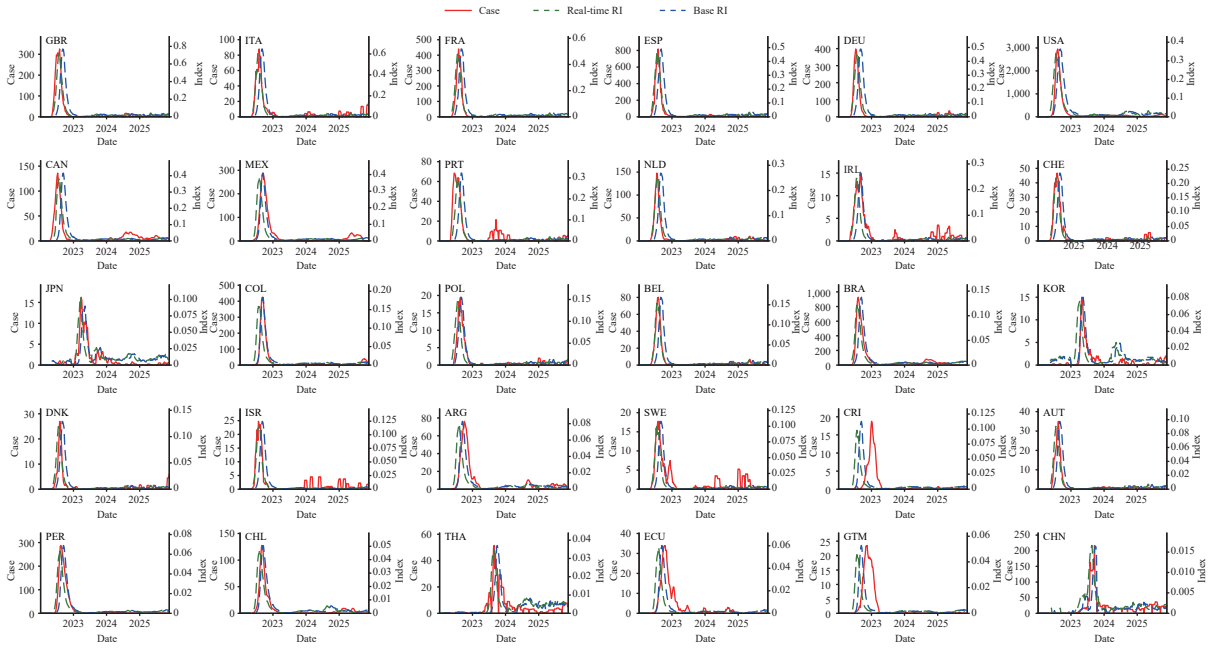


FIGURE 2. Country-specific time series of weekly cases and importation RI, 2022W01–2025W48.
Note: The weekly case series are shown as a red solid line (4-week moving average), the real-time RI as a green dashed line, and the baseline (base) RI as a blue dashed line. Cases are plotted on the left y-axis, while RI values are plotted on the right y-axis.
Abbreviation: RI=risk index; ARG=Argentina; AUT=Austria; BEL=Belgium; BRA=Brazil; CAN=Canada; CHE=Switzerland; CHL=Chile; CHN=China; COL=Colombia; CRI=Costa Rica; DEU=Germany; DNK=Denmark; ECU=Ecuador; ESP=Spain; FRA=France; GBR=United Kingdom; GTM=Guatemala; IRL=Ireland; ISR=Israel; ITA=Italy; JPN=Japan; KOR=Republic of Korea; MEX=Mexico; NLD=Netherlands; PER=Peru; POL=Poland; PRT=Portugal; SWE=Sweden; THA=Thailand; USA=United States.

The model demonstrates substantial early-warning value: compared with the baseline RI, the real-time RI provided an approximately four-week lead time. This advance warning can optimize public health resource allocation, reduce the economic impact of outbreaks, heighten public awareness, strengthen international cooperation, and inform evidence-based policy development. Further reductions in reporting delays would enable near-real-time situational awareness and extend the effective warning window (*12*), reinforcing the need for stronger international and multisectoral collaboration (*13–14*).

Beyond identifying when risk increases, the model helps pinpoint the sources of importation risk. For many countries, importation risk is concentrated among a set of major source countries, but the dominant sources differ across recipients. This heterogeneity supports tailored risk management strategies based on each country's connectivity profile (*15*). Our findings further suggest that even when overall global risk is low, localized or regional surges can rapidly reintroduce importation threats through highly connected travel nodes.

This study has several limitations. First, the RI relies on reported case counts and auxiliary indicators that may be affected by under-ascertainment, heterogeneous surveillance capacity, and reporting delays, potentially biasing estimates and limiting cross-country comparability. Second, the mobility component focuses exclusively on air travel and may underestimate risk in settings where land or sea crossings are substantial. Third, the linear approximation for $\lambda(t)$ may be inadequate during periods of rapid epidemic change; future work could explore piecewise-linear or nonlinear specifications to capture more complex trajectories. Fourth, equal weights were assumed for all regional risk index dimensions; future studies could leverage prior knowledge or empirical calibration data to optimize the weight vector and further enhance predictive accuracy.

Conflicts of interest: No conflicts of interest.

Acknowledgement: All individuals who generously shared their time and materials for this study.

Ethical Statement: All data used in this study are publicly available and do not involve personal

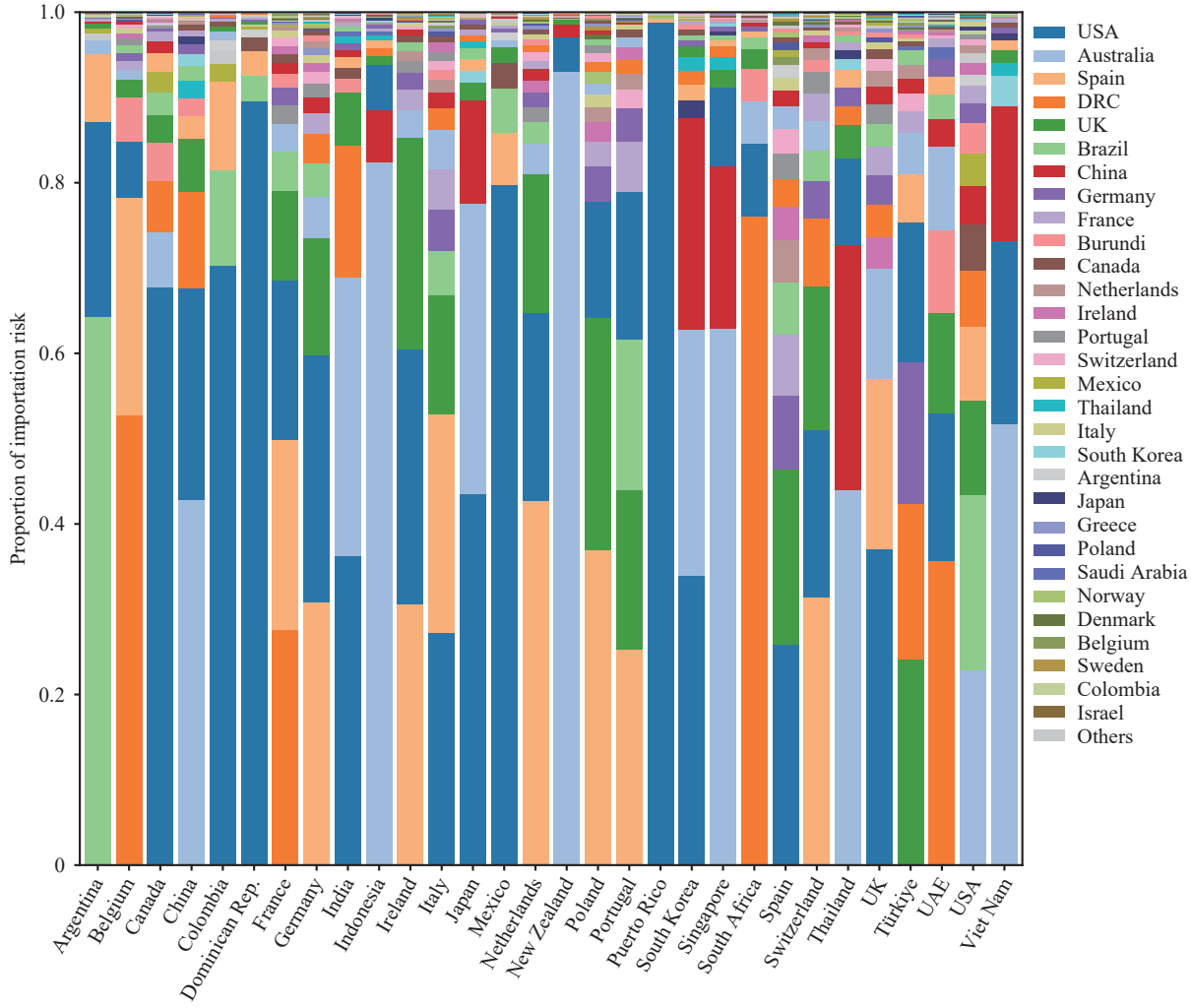


FIGURE 3. The proportion of importation risks for specific countries/regions.
Abbreviation: DRC=Democratic Republic of the Congo; UK=United Kingdom; USA=United States of America

information.

Funding: Supported by the Capital's Funds for Health Improvement and Research (2024-2G-3016), the National Key Research and Development Program of China (2023YFC2308701), the Key Research Projects of Beijing Natural Science Foundation-Haidian District Joint Fund (L242053), Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2022-ZHCH330-01), Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2023-I2M-3-011), Public Health Emergency Management Innovation Center, and State Key Laboratory of Respiratory Health and Multimorbidity.

doi: 10.46234/ccdcw2026.100

[#] Corresponding author: Zhongjie Li, lizhongjie@sph.pumc.edu.cn.

[1] School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [2] Beijing Office of Global Health, Beijing Center for Disease Prevention and Control, Beijing, China; [3] School of Public Health, Capital Medical University, Beijing, China; [4] Public Health Emergency Management Innovation Center of Beijing Higher Education Innovation Center for Philosophy and Social Sciences, Beijing, China; [5] State Key Laboratory of Respiratory Health and Multimorbidity, Beijing, China; [6] Key Laboratory of Pathogen Infection Prevention and Control (Peking Union Medical College), Ministry of Education, Beijing, China.
[&] Joint first authors.

Submitted: January 20, 2026
Accepted: May 08, 2026
Issued: May 15, 2026

REFERENCES

1. Laurenson-Schafer H, Sklenovská N, Hoxha A, Kerr SM, Ndumbi P, Fitzner J, et al. Description of the first global outbreak of mpox: an

analysis of global surveillance data. Lancet Glob Health 2023;11(7): e1012 – 23. https://doi.org/10.1016/S2214-109X(23)00198-5.
2. Mahase E. Mpox: US reports first clade Ib case as campaigners call for cheaper tests. BMJ 2024;387:q2576. https://doi.org/10.1136/bmj.q2576.
3. GOV. UK. Mpox outbreak: epidemiological overview, 11 December 2025. 2026. https://www.gov.uk/government/publications/monkeypox-outbreak-epidemiological-overview/mpox-outbreak-epidemiological-overview-11-december-2025?utm_source=chatgpt.com. [2026-1-10].
4. Au NH, Portillo MT, Marwah A, Thomas-Bachli A, Demarsh PA, Khan K, et al. Potential for monkeypox exportation from West and Central Africa through global travel networks. J Travel Med 2022;29 (6):taac072. https://doi.org/10.1093/jtm/taac072.
5. Gao S, Zeng Z, Xin Q, Yang MW, Feng XN, Liu XR, et al. Global transboundary transmission path and risk of Mpox revealed with Least Cost Path model. Int J Infect Dis 2024;146:107101. https://doi.org/10.1016/j.ijid.2024.107101.
6. Kinoshita R, Sassa M, Otake S, Yoshimatsu F, Shi S, Ueno R, et al. Impact of airline network on the global importation risk of mpox, 2022. Epidemiol Infect 2023;151:e60. https://doi.org/10.1017/S0950268823000456.
7. Huang QR, Sun YX, Jia MM, Jiang MY, Zhang T, Xu YS, et al. Risk assessment for cross-border transmission of multi-country Mpox outbreaks in 2022. J Infect Public Health 2023;16(4):618 – 25. https://doi.org/10.1016/j.jiph.2023.02.006.
8. WHO. 2022-24 Mpox (Monkeypox) Outbreak: Global Trends. 2024. https://reliefweb.int/report/world/2022-24-mpox-monkeypox-outbreak-global-trends. [2024-11-21].
9. Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021;5(4):529 – 38. https://doi.org/10.1038/s41562-021-01079-8.
10. Taylor L. Mpox in Africa: WHO and Africa CDC consider declaring public health emergency as cases spike. BMJ 2024;386:q1795. https://doi.org/10.1136/bmj.q1795.
11. Chan EH, Brewer TF, Madoff LC, Pollack MP, Sonricker AL, Keller M, et al. Global capacity for emerging infectious disease detection. Proc Natl Acad Sci USA 2010;107(50):21701 – 6. https://doi.org/10.1073/pnas.1006219107.
12. Du M, Sun HM, Zhang SM, Yuan J, Yan WX, Liu Q, et al. Global epidemiological features of human Monkeypox cases and their associations with social-economic level and international travel arrivals: a systematic review and ecological study. Int J Public Health 2023;68: 1605426. https://doi.org/10.3389/ijph.2023.1605426.
13. CDC. Infection prevention and control in healthcare settings. 2026. https://www.cdc.gov/mpox/hcp/infection-control/healthcare-settings.html. [2024-11-10].
14. Luo ST, Jiao KD, Zhang YH, Xu YT, Zhou JT, Huang SW, et al. Behavioral intention of receiving Monkeypox vaccination and undergoing Monkeypox testing and the associated factors among young men who have sex with men in China: large cross-sectional study. JMIR Public Health Surveill 2024;10(1):e47165. https://doi.org/10.2196/47165.
15. Prochazka M, Vinti P, Hoxha A, Seale A, Mozalevskis A, Lewis R, et al. Temporary adaptations to sexual behaviour during the mpox outbreak in 23 countries in Europe and the Americas: findings from a retrospective cross-sectional online survey. Lancet Infect Dis 2024;24 (12):1309 – 18. https://doi.org/10.1016/s1473-3099(24)00531-0.

Methods and Applications

Risk Assessment Method for Global Monkeypox Importation into China — May 2022–September 2025

Shihui Shan[1,&]; Yan Zhang[1,&]; Yanhe Wang[2,&]; Shenghong Lin[1]; Qiang Xu[1]; Chenlong Lyu[1]; Yao Tian[1]; Jun Ma[3]; Xinjing Zhao[1]; Yufeng Yang[1]; Yong Zhang[4,#]; Gang Dong[1,#]; Liqun Fang[1,#]

ABSTRACT

Introduction: In the context of globalization, the risk of cross-border transmission of infectious diseases is continuously increasing. There is an urgent need to develop a scientific, dynamic, and practical tool that captures regional heterogeneity in importation risk, thereby facilitating more targeted prevention and control measures at ports of entry.

Methods: We devised an extensible importation risk assessment framework by selecting key indicators through a Delphi expert consultation process and assigning their respective weights using the Analytic Hierarchy Process (AHP). Taking monkeypox (mpox) as a case study, we applied the framework to the global mpox epidemic spanning from May 2022 to September 2025 under three risk scenarios: Clade I, Clade II, and a non-clade-specific (aggregated) scenario. We then analyzed the temporal evolution of major source countries, the spatial distributions of provincial risk levels, and the association between clade-specific importation risk indices and domestically reported mpox cases across the study period.

Results: Our clade-stratified risk assessment framework demonstrates a high degree of consistency and reveals significant disparities between Clade I and Clade II in both international source contributions and provincial vulnerability. The importation risk associated with Clade I is predominantly influenced by Central African countries, particularly the Democratic Republic of the Congo and Uganda. In contrast, the risk related to Clade II is primarily associated with the United States and Southeast Asian nations, such as Thailand and the Philippines. Provincial importation risks generally adhere to a spatial pattern of "east-high, west-low": Shandong Province consistently exhibits a very high risk under Clade I, while Fujian Province maintains a very high risk under Clade II. Notably, the Clade II-specific risk index exhibited a statistically significant positive correlation with national mpox case counts (Spearman's ρ =0.50, P<0.05), whereas the non-clade-specific (aggregated) index did not (ρ =0.32, P=0.12).

Conclusion: This framework effectively captures the structural shift in the sources of mpox importation risk and the regional heterogeneity in risk levels across PLADs in China. It thereby provides a risk-driven decision-support framework for prioritizing port-of-entry screening and tailoring surveillance strategies to high-risk areas.

Globalization continues to raise the risk of cross-border infectious disease transmission (*1*). Since its discovery in 1958 (*2*), monkeypox (mpox) was long confined to Central and West Africa. It has two main clades: Clade I, prevalent in Central African countries like the Democratic Republic of the Congo (DR Congo), and Clade II, predominantly circulating in West African nations such as Nigeria and Ghana (*3–4*). In May 2022, Clade IIb sparked a global outbreak (*5*), prompting the World Health Organization (WHO) to declare a Public Health Emergency of International Concern (PHEIC) in July 2022 (*6*); that PHEIC ended in May 2023 (*7*). However, surging Clade I outbreaks in Central Africa led the WHO to declare a second PHEIC in August 2024 (*8*).

Existing frameworks for assessing infectious disease importation risk, such as those developed for coronavirus disease 2019 (COVID-19), dengue, and rabies, rely on aggregated case reports and travel flow data, not lineage- or clade-specific information (*9–12*). While useful for broad surveillance, they fail to capture how viral evolution (e.g., more transmissible or immune-evasive sub-lineages) changes importation risk. A few recent studies have begun incorporating variant data (*13*), but no standardized, quantitative, and operationally scalable framework for assessing the

importation risk of specific viral lineages or clades. To address this, we developed a modular, extensible framework that explicitly incorporates lineage-related characteristics into cross-border transmission risk estimation. Through disease-specific parameterization, this framework has potential applicability to various infectious diseases. Taking mpox as a case study, we applied it to three scenarios: Clade I, Clade II, and a non-clade-specific aggregated scenario. Leveraging multi-source authoritative data from May 2022 to September 2025, we conducted a systematic analysis on the spatiotemporal evolution of primary risk-source countries across three phases (first wave, inter-epidemic interval, second wave) and the dynamic distribution of risk levels across 31 provincial-level administrative divisions (PLADs) in the mainland of China. Additionally, correlations between clade-specific risk indices and reported case counts were evaluated.

METHODS

Selection of Risk-associated Indicators and Data Processing

Key risk-associated indicators were selected through the Delphi method and expert consultation (see the "Selection of Risk-Associated Indicators and Data Collection" section of the Supplementary Material, available at https://weekly.chinacdc.cn/). Twenty experts participated: 10 in Microbiology and Biosafety, 8 in Epidemiology, and 2 in Health Management. All held at least an associate senior professional title, had expertise in public health, infectious disease prevention and control, epidemiology, virology, or biosafety, and completed both Delphi rounds (demographics in Supplementary Table S1, available at https://weekly.chinacdc.cn/). System consistency was assessed using mean, authority coefficient, coefficient of variation (CV), and Kendall coordination coefficient (Kendall's W). The Analytic Hierarchy Process (AHP) was applied to determine the weights of the selected indicators.

We systematically collected data on risk-associated indicators from multiple sources, including: public health system capacity and socioeconomic factors (2019); Global Health Security Index (GHSI, 2021); air passenger flows (study period), epidemiological characteristics of infectious diseases (2024), and geographic proximity from 206 countries or territories to 31 PLADs in China (time-invariant). All quantitative indicators were standardized to 1−5 scores based on percentile rank for quantitative one and predefined criteria (disease characteristics, historical importation patterns, and geographic proximity) for qualitative ones. Full details including definitions, ranges, scoring methodologies, and data sources are provided in Supplementary Material (see "Selection of Risk-Associated Indicators and Data Collection" section).

Development of the Risk Assessment Framework

Considering the reliability of epidemic data from source countries and the representativeness of air-travel entry modes, an automated framework for assessing the importation risk of infectious diseases was developed, incorporating an incidence adjustment coefficient (S) and a connectivity index (C). To facilitate its practical application, the framework has been deployed on an infectious disease surveillance and early warning information system platform, enabling the automated calculations of weekly provincial risk indices from standardized data sources (details in the Supplementary Material). The risk index for month "a" (IT_a) is:

$$IT_a = [\ln(N) \times S + 1] \times C \quad (1)$$

Where N represents the number of newly reported cases in month "a", S indicates the adjustment coefficient of the incidence rate, derived from 14 weighted indicators: 5 on source-country public health capacity (F_1–F_5, such as universal health coverage, public health budget, health workforce, etc.) and 9 on disease biology (F_6–F_{14}, such as symptom typicality, fatality rate, incubation period, etc.) affecting both the pathogen's international transmission potential and clinical detectability. There is no potential overlap or double-counting of information between S and N. We apply the form $\ln(N)\times S$ to adjust observed case counts and better estimate cross-border importation risk. C quantifies the connectivity between the source country and 31 PLADs in the mainland of China, calculated from eight indicators (F_{15} to F_{22}) using their weights and scores. The details on the calculations of S and C are given in the "Development of the risk assessment framework" and "Details of the Framework Development" sections in the Supplementary Material. In primary analysis, months with zero cases (N=0) were excluded to avoid undefined $\ln(N)$; in sensitivity analysis, we used $\ln(N+1)$ to retain all observations.

Although demonstrated here with mpox, this modular framework can be adapted to other infectious diseases by re-scoring the S-module indicators

according to their epidemiological characteristics.

Risk Assessment and Correlation Analysis

This study formulated three risk scenarios based on clade-specific geographical distributions: Clade I (17 African countries with Clade I circulation), Clade II (African countries with Clade II circulation plus all non-African countries, considering that most non-African cases were Clade IIb), and non-clade-specific (aggregated). Using WHO's monthly global mpox case data (May 2022–September 2025) and international air passenger data (2022–2023), the monthly importation risk indices for 31 PLADs were calculated. The identical formula was used across all scenarios. Besides defining the source-country scope based on the known geographical distribution of each clade to produce separate risk estimate for each scenario, we also assigned clade-specific values to the adjustment coefficient (S) to reflect clinical differences, primarily in the "common symptoms" indicator (F_6): Clade I scored 1 (systemic symptoms like high fever); Clade II scored 3 (easily recognizable symptoms like low-grade fever and localized rash). This approach ensures that each risk index accurately reflects the specific epidemiological and connectivity characteristics of its clade or scenario.

The study covered three phases: the first wave (May 2022–May 2023), inter-epidemic interval (June 2023–July 2024), and the second wave (August 2024–September 2025). For Clade I, the temporal trends in risk contributions from all 17 African source countries were tracked across phases; for Clade II and non-clade-specific scenarios, the top 20 source countries by cumulative risk were presented. Provincial risk indices were categorized into five levels using quintiles. Confirmed mpox case data in China from September 2023 to September 2025 were obtained from China CDC. Validation used two complementary approaches: 1) Spearman correlation between reported case counts and monthly cumulative risk (Clade II and non-clade-specific scenarios); 2) qualitative spatial validation for Clade I, checking whether the PLADs with confirmed Clade I importations (Zhejiang and Shandong) (*14*–*15*) aligned with "extremely high" risk classifications. All analyses used SPSSAU (version 24.0; SPSSAU Project Team, Chongqing, China) and R (version 4.3.1; R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Risk-associated Indicators

Both consultation rounds achieved a 100% expert response rate. Based on expert feedback and group discussions, we developed an infectious disease importation risk assessment system comprising 2 primary, 4 secondary, and 22 tertiary indicators. The primary indicator X1 (epidemic situation in the source country) includes the epidemic detection capability (X1_1, F_1–F_5) and disease characteristics (X1_2, F_6–F_{14}). The primary indicator X2 (connectivity between source country and destination PLAD) includes the probability of importation from abroad (X2_1, F_{15}–F_{20}) and socioeconomic interactions (X2_2, F_{21}–F_{22}). The indicator importance rated by each expert are presented in detail in Supplementary Table S2 (available at https://weekly.chinacdc.cn/). The mean importance scores ranged from 3.650 to 4.950; coefficients of variation from 0.045 to 0.324; authority coefficients from 0.763 to 0.895. The Kendall's W was 0.443 (P<0.05), suggesting that the consultation results are reliable. The full statistics (mean, standard deviation, CV, authority coefficient) are summarized in Supplementary Table S3 (available at https://weekly.chinacdc.cn/). Details of the framework development can be found in the Supplementary Material. All secondary and tertiary indicator weights (Table 1) passed the consistency test [consistency ratio (CR)<0.1].

Risk Assessment

Clades show distinct imported risk source structures (Figure 1). In Clade I, dominant source countries shifted over time (Figure 1A). During the first wave, the Central African Republic, the DR Congo, and Sudan were primary contributors. The Central African Republic dominated in the early stage, then the DR Congo's contribution rose steadily. During the inter-epidemic interval, the DR Congo dominated almost exclusively, while contributions from other African countries decreased substantially. By the second wave, multi-center pattern emerged, with the DR Congo, Uganda, and Burundi led, while Kenya, Rwanda, Tanzania, and others also contributed significantly. For Clade II (Figure 1B), the United States remained a major risk source throughout all three phases. In the first wave, the primary contributors were the United States, Spain, and Nigeria. During the inter-epidemic interval, Nigeria's contribution declined substantially,

TABLE 1. Indicator weight distribution and testing.

First level indicator	Second level indicator	Weights	Third level indicator	Weights	Consistency check
Epidemic situation in the source country (X1)	Epidemic detection capability of source country (X1_1)	0.667	Effective universal health coverage in source country (F_1)	0.095	CI=0.043 CR=0.038 λ_{max}=5.172
			Public health budget of source country (F_2)	0.126	
			Health workforce in source country (F_3)	0.239	
			Pathogen diagnostic technology (F_4)	0.474	
			Global health security index (F_5)	0.069	
	Disease characteristics (X1_2)	0.333	Common symptom (F_6)	0.035	CI=0.049 CR=0.033 λ_{max}=9.388
			Case fatality rate (F_7)	0.135	
			Herd susceptibility (F_8)	0.253	
			Incubation period (F_9)	0.111	
			Infectiousness during incubation period (F_{10})	0.185	
			Communicable period (F_{11})	0.084	
			Classification of notifiable infectious disease (F_{12})	0.041	
			Classification of emerging infectious disease (F_{13})	0.050	
			Main routes of disease transmission (F_{14})	0.106	
Degree of connectivity between the source country and the PLAD in China (X2)	Probability of importation from abroad (X2_1)	0.750	Number of seats on flights from source country to the local area (F_{15})	0.355	CI=0.031 CR=0.024 λ_{max}=6.153
			Population migration index of the source country (F_{16})	0.117	
			History of similar disease importation from source country (F_{17})	0.158	
			Geographical location of the source country (F_{18})	0.062	
			Quantity of inbound overnight tourists received locally (F_{19})	0.084	
			Flight frequency from the source country to the local area (F_{20})	0.223	
	Socioeconomic interactions (X2_2)	0.250	Value of local imports and exports (F_{21})	0.333	CI=0 CR=0 λ_{max}=2
			Local international tourism revenue (F_{22})	0.667	

Note: The second level indicators under both primary indicators (X1 and X2) passed the consistency check (CI=0, CR=0, λ_{max}=2).
Abbreviation: CI=consistency index; CR=consistency ratio; λ_{max}=maximum eigenvalue of the judgment matrix.

and Thailand, Republic of Korea, and Vietnam emerged as primary sources. In the second wave, Nigeria rebounded, Thailand maintained high-risk, and the Philippines showed a notable rise. In the non-clade-specific scenario, the United States, Thailand, and the DR Congo were the primary risk sources across different phases, with shifting relative contributions over time (Supplementary Figure S1, available at https://weekly.chinacdc.cn/).

The risk of mpox importation into China exhibits significant spatiotemporal heterogeneity (Figure 2). For Clade I, cumulative risk increased gradually across three phases (range: 0–5 to 2–39) (Figure 2A). For Clade II, risk was higher during the first and second waves (range: 2–177 and 4–118, respectively), and dropped to 2–97 during the inter-epidemic interval

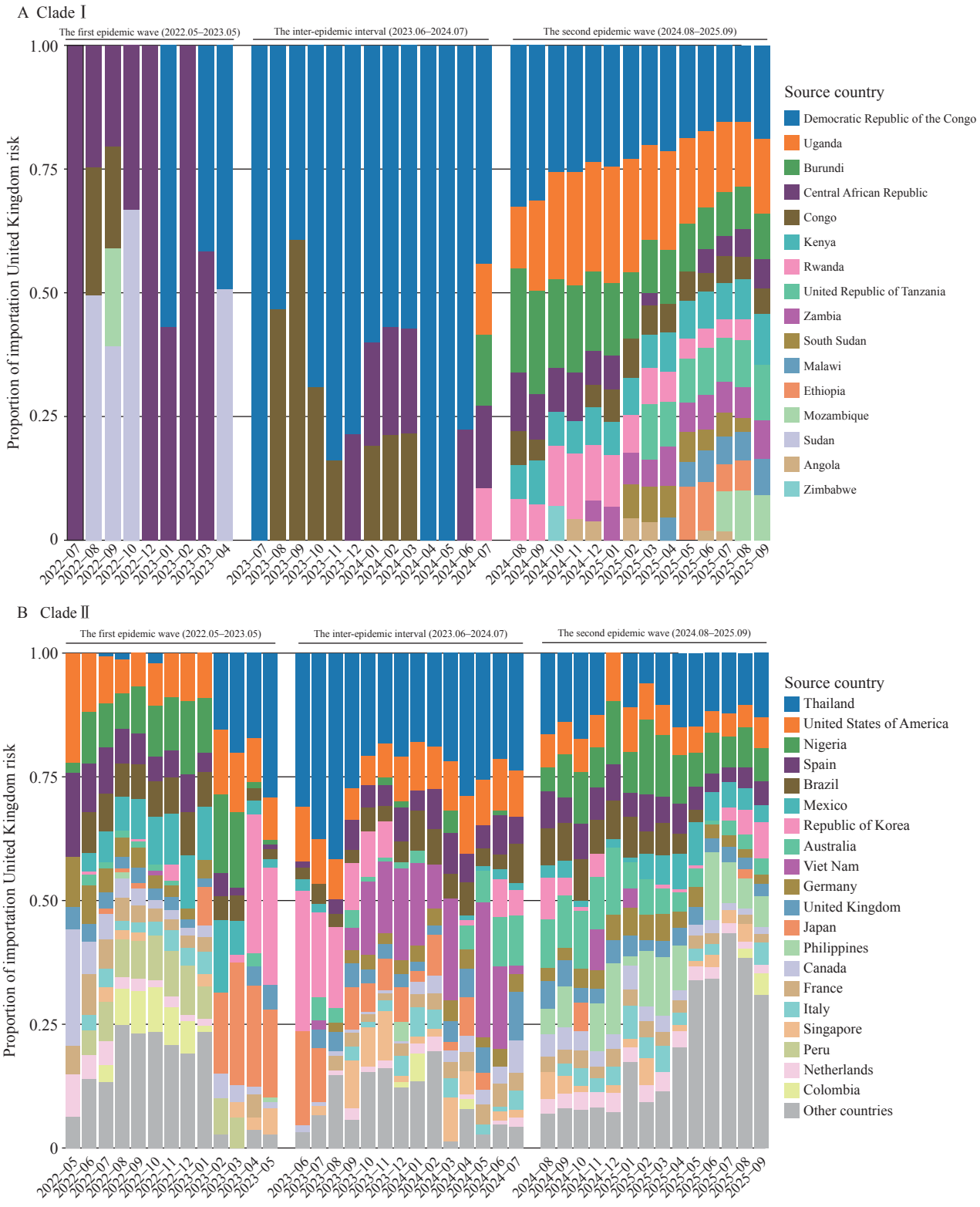


FIGURE 1. Clade-specific composition of source countries contributing to the risk of mpox importation across the three study phases. (A) Clade I: the proportions are calculated based on all identified source countries. (B) Clade II: the proportions are derived from the highest 10% of importation risk observations to highlight major source countries.
Abbreviation: mpox=monkeypox.

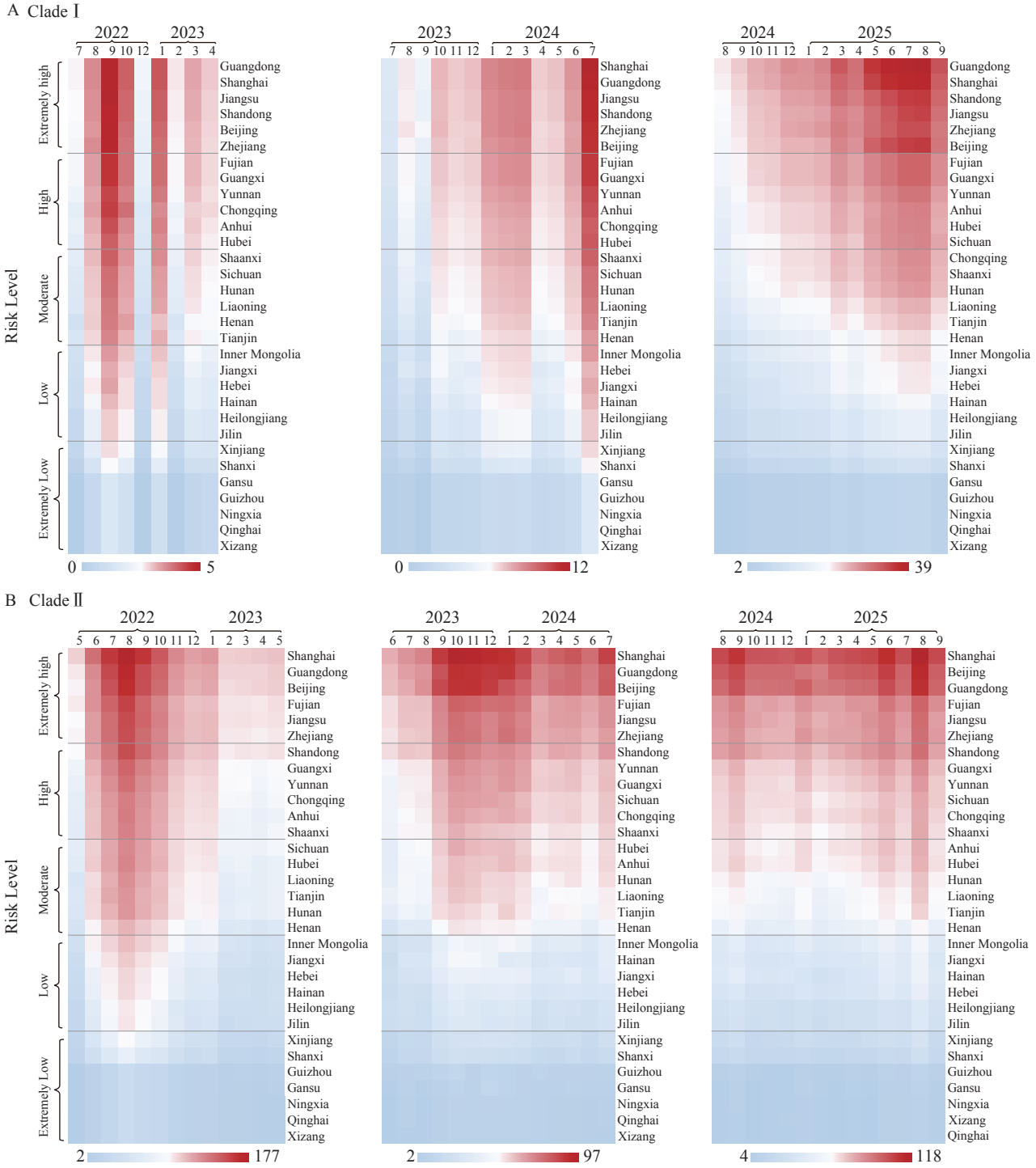


FIGURE 2. Importation risk of two mpox virus clades across 31 provincial-level administrative divisions (PLADs) in China. The heatmaps illustrate the estimated importation risks for (A) Clade I and (B) Clade II during the first epidemic wave (left panel), the inter-epidemic period (middle panel), and the second epidemic wave (right panel).
Note: Since the calculation of risk index is based on case data from source countries, and Clade I circulates in a limited number of countries, risk values could not be computed for certain months when no cases were reported in these source countries, resulting in temporal gaps in the Clade I time series.

(Figure 2B). Although overall provincial importation risk follows an "east-high, west-low" pattern, clade-stratified risks differ markedly in key PLADs. For instance, Shandong Province maintained extremely high-risk under Clade I but low-risk under Clade II, likely reflecting its role as a major maritime and air

gateway linking China with Africa. Conversely, Fujian Province consistently ranked as an extremely high-risk PLAD for Clade II, which is consistent with its extensive air links to Southeast Asia and other non-African countries that are the primary source regions of Clade II. Sichuan Province showed clade-dependent risk dynamics: high risk only during Clade I's second wave, but persistently high during both the inter-epidemic interval and second wave of Clade II. These patterns reveal that PLAD-level risk profiles arise from the interaction between each clade's geographical distribution and local international connectivity. As shown in Supplementary Table S4 (available athttps://weekly.chinacdc.cn/), the socio-geographic-mobility conditions (*A*) were the most important factors for nearly all country-PLAD pairs among Clade I source countries. In contrast, air travel conditions (*F*) dominated for high-risk PLADs (e.g., Shanghai, Guangdong, and Fujian) linked to major Clade II sources, particularly Thailand, the United States, the United Kingdom, and Australia. Provincial risk rankings in the non-clade-specific scenario generally resembled those in the Clade II scenario, but high-risk PLADs in Clade I (e.g., Shandong, Hubei) were significantly underestimated (Supplementary Figure S2, available at https://weekly.chinacdc.cn/).

Correlation Analysis

Time-series analysis revealed a significant positive correlation between national mpox case counts and the Clade II-specific importation risk index (ρ =0.50, P<0.05), but not for the non-clade-specific (aggregated) risk index (ρ =0.32, P=0.12) (Figure 3). Notably, Zhejiang and Shandong provinces, both reporting primary Clade I imported cases, were consistently assessed as having extremely high Clade I importation risk throughout the study period. Sensitivity analysis using ln(N+1) instead of ln(N) yielded near-identical risk estimates (Clade II: ρ =0.993; Clade I: ρ =0.995; both P<0.001), indicating the robustness of this risk assessment framework.

DISCUSSION

This study introduces the two predominant mpox virus clades into the importation risk assessment framework, revealing distinct risk patterns between Clade I and Clade II. For Clade I, risk sources shifted from the Central African Republic (first wave) to a multi-center pattern in the second wave, led by the DR Congo, Uganda, and Burundi. For Clade II, the United States remained a consistently major source

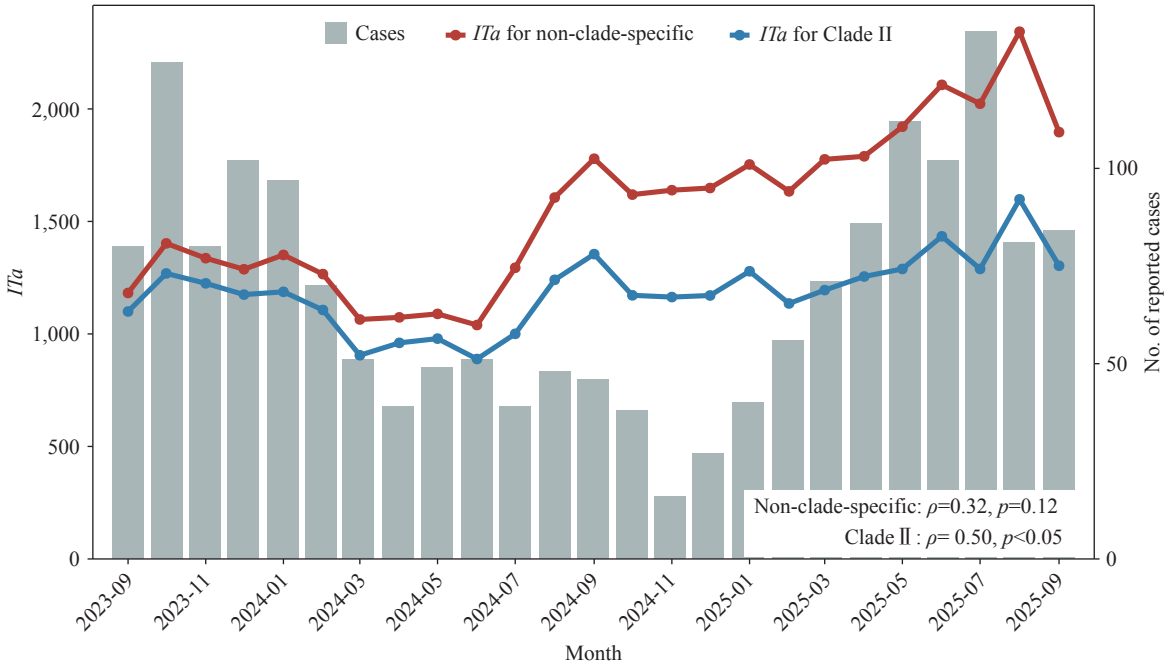


FIGURE 3. Temporal trends in the importation risk index and the number of reported mpox cases at the national level in China from September 2023 to September 2025.
Note: ρ denotes the Spearman correlation coefficient between IT_a and the reported number of mpox cases.

throughout all phases, but during the inter-epidemic interval, risk sources shifted toward Asian countries, particularly Thailand, Republic of Korea, and Vietnam. These dynamics align with the global epidemiological patterns documented by Cevik et al. They reported intensified Clade I outbreaks in Central Africa and the continued low-level Clade IIb circulation in non-endemic regions (*16*). The observed shift of Clade II risk sources towards Southeast Asia during the inter-epidemic interval is consistent with the regional surveillance data, which indicates sustained transmission in Thailand and neighboring countries (*17*).

Shandong Province had consistently very high importation risk under Clade I throughout all three phases, but low risk under Clade II. Fujian Province exhibited the opposite pattern. Non-clade-specific analysis substantially underestimated these differences, highlighting fundamental differences in geographical distribution and transmission dynamics between the two clades, as emphasized by Srivastava et al., and underscoring the need for clade-specific surveillance (*3*).

Clade II's risk index exhibited a statistically significant, moderate positive correlation with reported case counts (ρ =0.50, P<0.05), whereas the non-clade-specific aggregated index did not (ρ =0.32, P=0.12). This indicates that almost all mpox cases in China during 2023–2025 belonged to Clade IIb, except for a few in Shandong, Zhejiang, and Beijing (*14–15*,*18*). This is consistent with Clade IIb's global dominance outside Africa since 2022 (*19*). Thus, overseas importation risk is a crucial factor contributing to the observed domestic epidemic pattern. Factor decomposition analysis (Supplementary Table S4) further revealed that the spatial heterogeneity in provincial risk is shaped by different factors across clades: for Clade I, socio-geographic-mobility conditions (*A*) dominate due to minimal direct air links from African source countries; for Clade II, air travel conditions (*F*) serve as the most important factor for high-risk PLADs with direct flights from major source countries (e.g., Thailand, the United States, and Australia), while *A* remains dominant where such flights are limited. Temporal variation in risk ranking stems mainly from changes in source-country epidemic magnitude (*N*), accounting for the evolving contributions of different countries across the three study phases (Figure 1). In addition to the quantitative correlation, the spatial risk classifications of the framework are further supported by real-world events. During the study period, Zhejiang and Shandong provinces had documented Clade I importation cases (*14–15*), which were consistently classified in the highest risk quintile by the framework throughout all three phases, providing preliminary directional support for the framework's discriminative capacity under data-scarce conditions.

In comparison with existing frameworks, this study offers three potential improvements over previous approaches. Travel-volume-based models applied to COVID-19 and dengue estimate risk solely from air flows and source-country incidence, without considering detection capability or disease characteristics (*9–10*,*20*). AHP-based frameworks assess regional vulnerability in China, but concentrate on domestic transmission, not cross-border importation (*21*). The current framework integrates both the epidemic characteristics of the source country and the connectivity of the destination PLAD into a unified index, introduces clade-stratified differentiation to capture the divergent geographical footprints of pathogens, and provides provincial risk rankings across all 31 PLADs in China to guide the prioritization of resources at ports of entry. These findings have significant implications for public health. First, the framework supports regionally tailored surveillance strategies with scientific evidence. Second, the dynamic monitoring of the contributions from risk source countries allows for the timely adjustment to border screening and control measures as global outbreaks evolve. Third, the recent Clade Ib cluster in Yiwu City of Zhejiang Province demonstrates that international trade hubs remain vulnerable to rapid local transmission, highlighting the necessity for enhanced inbound surveillance targeting travelers from high-prevalence regions (*15*). China's comprehensive intervention framework, including early surveillance, mandatory case reporting, and timely treatment, has proven to be effective in containing imported cases and preventing community transmission (*22*).

This study has two main limitations. First, the China CDC reported imported and locally transmitted cases together, not separately. Since our risk index estimates only importation risk, secondary cases likely weakened the observed correlation, particularly during the second wave, when local clusters were more prevalent. This may partly account for the moderate magnitude of the Clade II correlation coefficient (ρ =0.50), despite its statistical significance. Nevertheless, the association indicates that the risk index captures meaningful spatiotemporal variability in

the early importation dynamics, which holds practical value for identifying PLADs that require intensified border screening and rapid-response preparedness. Second, in order to approximate the global distribution of Clade II, it was assumed that all mpox cases reported to WHO from non-African countries belonged to Clade II. While this assumption largely aligns with current epidemiological evidence, given that Clade IIb has dominated transmission outside Africa since 2022, it may overlook rare, undetected Clade I cases, potentially introducing misclassification bias. Future research could integrate multimodal traffic flow, social media sentiment, and environmental wastewater monitoring data, and extend the scope to other imported diseases.

Conflicts of interest: No conflicts of interest.

Acknowledgements: All individuals who participated in the research and all organizations that provided data.

Funding: Supported by Beijing Natural Science Funds (L242053).

doi: 10.46234/ccdcw2026.101

[#] Corresponding authors: Yong Zhang, zhangyong@bnu.edu.cn; Gang Dong, dongg999@sina.com; Liqun Fang, fanglq@bmi.ac.cn.

[1] State Key Laboratory of Pathogen and Biosecurity, Academy of Military Medical Sciences, Beijing, China; [2] The 968th Hospital of Joint Logistics Support Force of PLA, Jinzhou City, Liaoning Province, China; [3] School of Public Health, the Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Medical University, Guiyang City, Guizhou Province, China; [4] School of Mathematical Sciences, Beijing Normal University, Beijing, China.
[&] Joint first authors.

Submitted: January 19, 2026
Accepted: May 01, 2026
Issued: May 15, 2026

REFERENCES

1. Frenk J, Gómez-Dantés O, Knaul FM. Globalization and infectious diseases. Infect Dis Clin North Am 2011;25(3):593 – 9. https://doi.org/10.1016/j.idc.2011.05.003.
2. Parker S, Buller RM. A review of experimental and natural infections of animals with monkeypox virus between 1958 and 2012. Future Virol 2013;8(2):129 – 57. https://doi.org/10.2217/fvl.12.130.
3. Srivastava S, Sharma D, Sridhar SB, Kumar S, Rao GSNK, Budha RR, et al. Comparative analysis of mpox clades: epidemiology, transmission dynamics, and detection strategies. BMC Infect Dis 2025;25(1):1290. https://doi.org/10.1186/s12879-025-11784-8.
4. Kinganda-Lusamaki E, Amuri-Aziza A, Fernandez-Nuñez N, Makangara-Cigolo JC, Pratt C, Vakaniaki EH, et al. Clade I mpox virus genomic diversity in the Democratic Republic of the Congo, 2018-2024: predominance of zoonotic transmission. Cell 2025;188(1):4 – 14.e6. https://doi.org/10.1016/j.cell.2024.10.017.
5. Vakaniaki EH, Kacita C, Kinganda-Lusamaki E, O'Toole Á, Wawina-Bokalanga T, Mukadi-Bamuleka D, et al. Sustained human outbreak of a new MPXV clade I lineage in eastern Democratic Republic of the Congo. Nat Med 2024;30(10):2791 – 5. https://doi.org/10.1038/s41591-024-03130-3.
6. World Health Organization. Second meeting of the International Health Regulations (2005) (IHR) Emergency Committee regarding the multi-country outbreak of monkeypox. 2022. https://www.who.int/news/item/23-07-2022-second-meeting-of-the-international-health-regulations-(2005)-(ihr)-emergency-committee-regarding-the-multi-country-outbreak-of-monkeypox. [2026-5-12].
7. World Health Organization. Fifth meeting of the international health regulations (2005) (IHR) emergency committee on the multi-country outbreak of mpox (monkeypox). 2023. https://www.who.int/news/item/11-05-2023-fifth-meeting-of-the-international-health-regulations-%282005%29-%28ihr%29-emergency-committee-on-the-multi-country-outbreak-of-monkeypox-%28mpox%29. [2026-5-12].
8. World Health Organization. WHO director-general declares mpox outbreak a public health emergency of international concern. 2024. https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern. [2026-5-12].
9. Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boëlle PY, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. Lancet 2020;395(10227):871 – 7. https://doi.org/10.1016/S0140-6736(20)30411-6.
10. Poongavanan J, Lourenço J, Tsui JLH, Colizza V, Ramphal Y, Baxter C, et al. Dengue virus importation risks in Africa: a modelling study. Lancet Planet Health 2024;8(12):e1043 – 54. https://doi.org/10.1016/S2542-5196(24)00272-9.
11. Min KD, Kim SY, Cho YY, Kim S, Yeom JS. Potential applicability of the importation risk index for predicting the risk of rarely imported infectious diseases. BMC Public Health 2023;23(1):1776. https://doi.org/10.1186/s12889-023-16380-6.
12. Milwid RM, Gabriele-Rivet V, Ogden NH, Turgeon P, Fazil A, London D, et al. A methodology for estimating SARS-CoV-2 importation risk by air travel into Canada between July and November 2021. BMC Public Health 2024;24(1):1088. https://doi.org/10.1186/s12889-024-18563-1.
13. Asakura TR, Jung SM, Jin SH, Hu G, Endo A, Dickens BL. Characterising global risk profiles of Mpox clade Ib importation. J Travel Med 2024;31(8):taae136. https://doi.org/10.1093/jtm/taae136.
14. Yin CH, Li Y, Duan Q, Xu DL, Li CYX, Liu T, et al. Genomic and phenotypic insights into the first imported monkeypox virus clade Ia isolate in China, 2025. Front Public Health 2025;13:1618022. https://doi.org/10.3389/fpubh.2025.1618022.
15. Long ZP, Zhu LB, Cai J, Ji HR, Ye Z, Shi XG, et al. The cluster of mpox (Clade Ib) infections — Yiwu City, Zhejiang Province, China, July–August 2025. China CDC Wkly 2025;7(41):1308 – 13. https://doi.org/10.46234/ccdcw2025.221.
16. Cevik M, Tomori O, Mbala P, Scagliarini A, Petersen E, Low N, et al. The 2023-2024 multi-source mpox outbreaks of Clade I MPXV in sub-Saharan Africa: alarm bell for Africa and the World. Int J Infect Dis 2024;146:107159. https://doi.org/10.1016/j.ijid.2024.107159.
17. Akhavan B, Ahmed AB, To W, Alkhouri S, Batra K. From forest to frontline: a comprehensive review of mpox's global leap and viral evolution (2022-2024). One Health 2025;21:101232. https://doi.org/10.1016/j.onehlt.2025.101232.
18. Bian HL, Wang XW, Li S, Chen CZ, Chen XM, Zhu JY, et al. Epidemiological investigation and disposal of the first confirmed infection case of mpox virus Ib in Beijing. Dis Surveill 2025;40(3):341 – 5. https://doi.org/10.3784/jbjc.202502110079.

19. Ndembi N, Folayan MO, Komakech A, Mercy K, Tessema S, Mbala-Kingebeni P, et al. Evolving epidemiology of mpox in Africa in 2024. N Engl J Med 2025;392(7):666 – 76. https://doi.org/10.1056/NEJMoa2411368.
20. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 2020;368(6489):395 – 400. https://doi.org/10.1126/science.aba9757.
21. Gao ZK, Jiang YT, He JY, Wu JP, Xu J, Christakos G. An AHP-based regional COVID-19 vulnerability model and its application in China. Model Earth Syst Environ 2022;8(2):2525 – 38. https://doi.org/10.1007/s40808-021-01244-y.
22. Shang WJ, Yan J, Zhang HB, Liang WN, Liu J, Liu M. Mpox containment strategies in China, 2022-2023: from preparedness to legal management. BMJ Glob Health 2025;10(8):e019279. https://doi.org/10.1136/bmjgh-2025-019279.

SUPPLEMENTARY MATERIALS

Selection of Risk-Associated Indicators and Data Collection

Details of the indicator selection: In the first round of the questionnaire, the total weight of the secondary indicators under each primary indicator and the total weight of the tertiary indicators under each secondary indicator were 1,000. Experts assigned weights to the secondary and tertiary indicators in the system. Indicators that were assigned zero weight by two or more experts, as well as meaningless indicators, were not included in the second round of the questionnaire. Indicators considered meaningful by the experts (which were both accessible and quantifiable) were added on this basis. In the second round of the questionnaire, experts assessed the importance, basis of judgment, and familiarity of all indicators. Any suggestion to delete an indicator was discussed in depth to decide whether to retain it.

Specific details of data collection are as follows: Data on the public health systems of various countries and territories were gathered from the Global Burden of Disease (GBD) database and the Global Health Security Index (GHSI) databases. These data encompassed the following indicators: effective universal health coverage (F_1), public health budget (F_2), and health workforce (F_3, sourced from the GBD database), as well as the Global Health Security Index (F_5, sourced from the GHSI database). F_1, F_2, F_3, and F_5 were scored using an inverse percentile rank: countries in the 81–100th percentile received a score of 1, and those in the 0–20th percentile received a score of 5, reflecting that stronger public health capacity leads to better case detection and lower underreporting risk. This system incorporates indicator data from 207 countries and territories. For the few territories with missing data, a score of 3 was used as a supplement. Data were missing for 3 countries in the 4 indicators of effective universal health coverage in the source country, public health budget in the source country, health workforce in the source country, and population migration index in the source country, and for 12 countries in the Global Health Security Index, and a score of 3 was used in this study as a proxy score for the missing values. This score represents 'neutral' or 'uncertain' status to avoid bias in scoring due to missing data and to ensure comparability of scores across countries.

Population movement data are sourced from the Official Aviation Guide (OAG) and the GBD database. These datasets encompass several key indicators, including the number of seats on flights from the source country to the destination region (F_{15}), flight frequency from the source country to the destination region (F_{20}, both sourced from OAG database), and the population migration index of the source country (F_{16}, sourced from the GBD database).

SUPPLEMENTARY TABLE S1. Experts' demographic information.

Characteristics	Demographics	Count
Gender	Male	16
	Female	4
Age (years)	0–44	8
	45–54	10
	55–64	2
Primary responsibility	Teaching & Research	16
	Clinical work	1
	Administrative management	1
	Other	2
Title of the job	Professor	13
	Associate professor	7
Education level	Doctoral/PhD	19
	Master	1
Professional field	Microbiology and Biosafety	10
	Epidemiology	8
	Health management	2

SUPPLEMENTARY TABLE S2. The importance scoring by each expert for each indicator.

Indicator	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
X1_1	4	5	5	4	4	5	5	4	5	5	5	5	4	5	5	5	5	5	5	5
X1_2	4	5	5	5	5	5	4	4	5	5	5	4	4	4	5	5	5	5	5	4
X2_1	3	5	5	5	5	5	5	5	5	5	5	5	4	4	5	3	5	5	5	4
X2_2	4	5	5	4	4	3	3	4	4	5	4	4	5	3	5	4	4	4	5	5
F_1	2	5	5	4	4	4	3	4	5	5	4	4	4	3	5	4	3	4	5	3
F_2	3	4	5	5	4	3	4	3	4	5	3	4	4	3	5	4	4	5	5	4
F_3	4	5	5	5	4	4	4	3	5	5	4	4	4	3	5	4	5	5	5	4
F_4	5	5	5	4	5	5	5	5	5	5	5	5	4	4	5	5	5	5	5	4
F_5	2	4	5	3	5	4	5	3	4	3	4	3	4	4	5	3	4	5	5	3
F_6	2	5	5	5	4	4	4	4	5	5	3	3	4	3	5	3	4	5	5	4
F_7	5	5	5	5	4	5	5	5	5	5	5	5	5	4	5	2	5	5	5	3
F_8	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
F_9	3	5	5	5	5	4	5	5	5	5	5	5	4	4	5	4	4	5	5	4
F_{10}	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	3
F_{11}	3	5	5	5	5	5	4	5	5	5	5	4	4	4	5	5	4	5	5	3
F_{12}	2	4	5	4	4	4	3	4	4	5	4	4	4	5	5	4	4	5	5	4
F_{13}	2	5	5	4	4	4	3	4	4	5	4	4	4	5	5	4	4	5	5	5
F_{14}	2	5	5	5	5	5	5	5	5	5	5	4	5	4	5	4	4	5	5	4
F_{15}	3	5	5	5	4	3	4	4	4	5	4	5	4	4	5	4	5	5	5	4
F_{16}	4	5	5	5	5	4	4	5	5	5	5	4	4	4	5	4	5	5	5	3
F_{17}	4	5	5	4	4	3	4	3	5	5	4	3	4	3	5	3	4	5	5	4
F_{18}	3	5	5	4	5	3	3	3	5	5	4	4	4	5	5	2	5	5	5	3
F_{19}	2	4	5	3	4	5	4	2	4	4	4	3	4	3	5	3	3	5	5	5
F_{20}	2	5	5	5	3	4	3	3	4	5	4	3	4	4	5	4	4	4	5	4
F_{21}	3	4	5	4	4	1	5	3	5	5	3	3	3	2	5	2	4	4	5	3
F_{22}	3	5	5	3	4	3	5	3	5	5	4	4	4	3	5	2	3	4	5	2

Note: Importance scores were defined as follows: 1 indicates very unimportant; 2 indicates unimportant; 3 indicates moderately importance; 4 indicates important; 5 indicates very important. Experts are denoted by letters A–T.

Socioeconomic data were obtained from the China Statistical Yearbooks Database (CSYD), and indicators included the number of inbound travelers to PLADs in China (F_{19}), value of local imports and exports (F_{21}), and local international tourism revenue (F_{22}). The scoring methodology for population movement and socioeconomic indicators (F_{15}, F_{16}, F_{19}, F_{20}, F_{21}, F_{22}) followed a positive percentile rank: countries or PLADs in the 0–20th percentile receive a score of 1, and those in the 81–100th percentile received a score of 5.

In the risk assessment framework, a historical disease importation indicator (F_{17}) was developed for each source country-disease-PLAD combination, based on the recorded number of similar infectious disease cases imported from source countries to destination PLADs between 2014 and 2015, as reported by the General Administration of Customs of China (GACC) and China CDC. The scoring rules are based on the frequency and magnitude of historical imports; if there are no records of imports for certain diseases, imports of diseases with the same transmission routes are taken into account, and additionally importation events with more than 10 cases in a two-year period are defined as "large-scale", otherwise they are considered as "small-scale". Specifically: a score of 1 is assigned if there is no record of similar disease importation; a score of 2 is given for small-scale importation of a similar disease but no record of this disease; a score of 3 for multiple or large-scale imports of a similar disease but no record of this disease; and a score of 4 for small-scale import record of the disease; records of multiple or large-scale importation of the disease were 5 points.

Furthermore, disease-related characteristic data were obtained for each recorded infectious disease from

SUPPLEMENTARY TABLE S3. Mean value, standard deviation, coefficient of variation, and authority coefficient of importance scores for each indicator.

Indicator	Mean value	Standard deviation	Coefficient of variation	Authority coefficient
X1_1	4.750	0.444	0.094	0.884
X1_2	4.650	0.489	0.105	0.881
X2_1	4.650	0.671	0.144	0.870
X2_2	4.200	0.696	0.166	0.816
F_1	4.000	0.858	0.215	0.823
F_2	4.050	0.759	0.187	0.816
F_3	4.350	0.671	0.154	0.823
F_4	4.800	0.410	0.085	0.886
F_5	3.900	0.912	0.234	0.800
F_6	4.100	0.912	0.222	0.853
F_7	4.650	0.813	0.175	0.894
F_8	4.950	0.224	0.045	0.895
F_9	4.600	0.598	0.130	0.893
F_{10}	4.850	0.489	0.101	0.891
F_{11}	4.550	0.686	0.151	0.876
F_{12}	4.150	0.745	0.180	0.864
F_{13}	4.250	0.786	0.185	0.868
F_{14}	4.600	0.754	0.164	0.894
F_{15}	4.550	0.605	0.133	0.834
F_{16}	4.100	0.788	0.192	0.803
F_{17}	4.150	0.988	0.238	0.843
F_{18}	3.850	0.988	0.257	0.813
F_{19}	4.000	0.858	0.215	0.788
F_{20}	4.350	0.670	0.150	0.820
F_{21}	3.650	1.182	0.324	0.763
F_{22}	3.850	1.040	0.270	0.781

authoritative sources including the Global Infectious Diseases and Epidemiology Network (GIDEON), WHO, and China CDC. These data encompassed the following indicators: pathogen diagnostic technology, common symptom, case fatality rate, herd susceptibility, incubation period, infectiousness during incubation period, communicable period, and the main routes of disease transmission, classification of infectious diseases. Scoring criteria for these indicators were predefined as follows. Pathogen diagnostic technology (F_4): score 1 if mature diagnostic techniques are available; score 2 if established but unverified; score 3 if techniques exist for similar pathogens only; score 4 if no techniques exist for similar pathogens; score 5 if the etiology is unknown. Common symptom (F_6): score 1 if high fever and widespread rash are present; score 3 if low-grade fever and localized rash are present; score 5 if fever and rash are absent. Case fatality rate (F_7): score 5 if <1%; score 3.75 if 1%–10%; score 2.5 if 10%–30%; score 1.25 if >30%. Herd susceptibility (F_8): score 1 if a limited population with high herd immunity; score 2 if a widespread population with high herd immunity; score 3 if generally susceptible with high herd immunity, or a limited population with low herd immunity; score 4 if a widespread population with low herd immunity; score 5 if generally susceptible with low herd immunity. Incubation period (F_9) and communicable period (F_{11}): score 1.25 if <3 days; score 2.5 if 3–7 days; score 3.75 if 7–14 days; score 5 if >14 days or uncertain. Infectiousness during incubation period (F_{10}): score 1 if no infectivity; score 3 if weak infectivity; score 5 if strong infectivity or uncertain. Notifiable disease classification (F_{12}): score 1 if Class A; score 3 if Class B; score 5 if Class C or other. Emerging disease classification (F_{13}): score 1 if not newly emerging in China; score 3 if newly emerging in

SUPPLEMENTARY TABLE S4. The most important factor driving these spatial and temporal variations of the risk index of the cross-border transmission of mpox (PLADs with top five cumulative risk), stratified by clades (May 2022–September 2025).

Clade	Source country	Cumulative *ITa*	The most important factor/PLAD1	The most important factor/PLAD2	The most important factor/PLAD3	The most important factor/PLAD4	The most important factor/PLAD5
Clade I	Democratic Republic of the Congo	2,329.7	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)	Beijing (A)
	Uganda	1,203.8	Shanghai (A)	Zhejiang (A)	Hubei (A)	Guangdong (A)	Shandong (A)
	Burundi	842.9	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)	Beijing (A)
	Central African Republic	660	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)	Beijing (A)
	Congo	526.6	Zhejiang (A)	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)
	Kenya	497	Guangdong (F)	Sichuan (F)	Shandong (A)	Hunan (A)	Shanghai (A)
	Rwanda	445.9	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)	Beijing (A)
	United Republic of Tanzania	381.9	Guangdong (A)	Shanghai (A)	Shandong (A)	Beijing (A)	Jiangsu (A)
	Zambia	328.6	Shandong (A)	Zhejiang (A)	Shanghai (A)	Guangdong (A)	Jiangsu (A)
	South Sudan	201.7	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)	Beijing (A)
Clade II	Thailand	2,405.1	Shanghai (F)	Beijing (F)	Zhejiang (F)	Yunnan (F)	Fujian (F)
	United States of America	2,042.5	Shanghai (F)	Guangdong (F)	Fujian (F)	Beijing (A)	Jiangsu (A)
	Nigeria	1,985	Beijing (A)	Guangdong (A)	Shanghai (A)	Shandong (A)	Jiangsu (A)
	Spain	1,916.5	Beijing (A)	Shanghai (A)	Zhejiang (A)	Jiangsu (A)	Chongqing (A)
	Mexico	1,790	Beijing (A)	Zhejiang (A)	Shanghai (A)	Guangdong (A)	Shandong (A)
	Brazil	1,696	Shanghai (A)	Jiangsu (A)	Chongqing (A)	Guangdong (A)	Shandong (A)
	Colombia	1,361.1	Shanghai (A)	Guangdong (A)	Shandong (A)	Jiangsu (A)	Beijing (A)
	The United Kingdom	1,357.3	Shanghai (F)	Beijing (F)	Guangdong (A)	Shandong (A)	Shaanxi (A)
	Canada	1,301.9	Shanghai (F)	Guangdong (A)	Fujian (A)	Beijing (A)	Shandong (A)
	Australia	1,252.2	Guangdong (F)	Shanghai (F)	Fujian (F)	Beijing (A)	Jiangsu (A)

Note: The IT_a is a product of epidemic magnitude ($\ln(N+1)\cdot S$) and the composite connectivity score C ($C=A+B+F$). Temporal variations in the risk contributed by each source country are primarily driven by fluctuations in N over time. Spatial variations across PLADs within the same source country are determined by C, as $\ln(N+1)\cdot S$ is country-specific and does not vary across provinces. The most important factor shown in each cell refers to the largest component of C for that country–PLAD pair. B did not appear as the most important factor for any pair due to its limited value range (0–0.25).

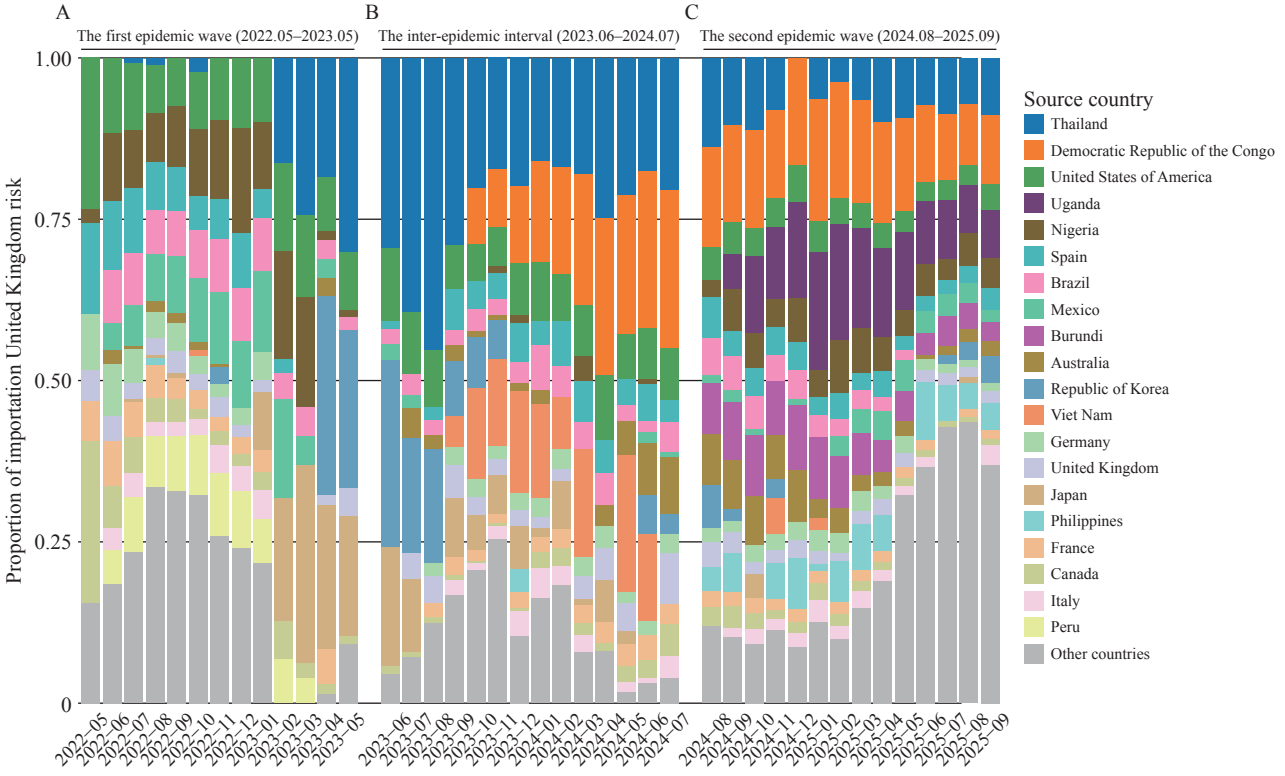
Abbreviation: mpox=monkeypox; IT_a=risk index of cross-border transmission of infectious diseases in month a; N=number of newly reported cases; S=adjustment coefficient of incidence rate; C=composite connectivity score; A=normalized socio-geographic and population mobility index of the destination PLAD; B=normalized historical importation record of similar infectious diseases from the source country; F=normalized air travel connectivity between the source country and the destination PLAD.

China; score 5 if newly emerging globally. Main transmission routes (F_{14}): score 1 for surface contact-borne; score 2 for bloodborne/sexually transmitted; score 3 for vector-borne; score 4 for gastrointestinal; score 5 for respiratory.

Geographic data were obtained from the Ministry of Natural Resources of the People's Republic of China and include indicators of the geographical location of the source country (F_{18}). Data were processed on a scale of 1 to 5 depending on the geographical proximity of the source country to China, far away from China (1.25); not sharing a land border with China but facing each other across the sea (2.5); not sharing a land border with China but coexisting on the same landmass (3.75); sharing a land border with China (5).

Development of the Risk Assessment Framework

Incidence rate correction coefficient (S): Given the potential for underreporting of infectious disease cases in the source country, it is deemed necessary to correct the incidence rate using the incidence rate correction coefficient



SUPPLEMENTARY FIGURE S1. Non-clade-specific of source countries contributing to mpox importation risk across the three study phases.
Abbreviation: mpox=monkeypox.

(S). This coefficient is determined by considering two factors: the epidemic detection capability of source country and the disease characteristics. This coefficient is calculated using 14 indicators, including effective universal health coverage in source country (F_1), public health budget of source country (F_2), and health workforce in source country (F_3) among others up to F_{14}. The specific calculation formula is as follows:

$$S = \frac{\sum_{i=1}^{14} a_i \times F_i \times b_i}{\ln 100} \quad (2)$$

Where a_i denotes the weight coefficient of the secondary indicators, b_i denotes the weight coefficient of the tertiary indicators, and F_i represents the score of the tertiary indicators (on a scale from 1 to 5).

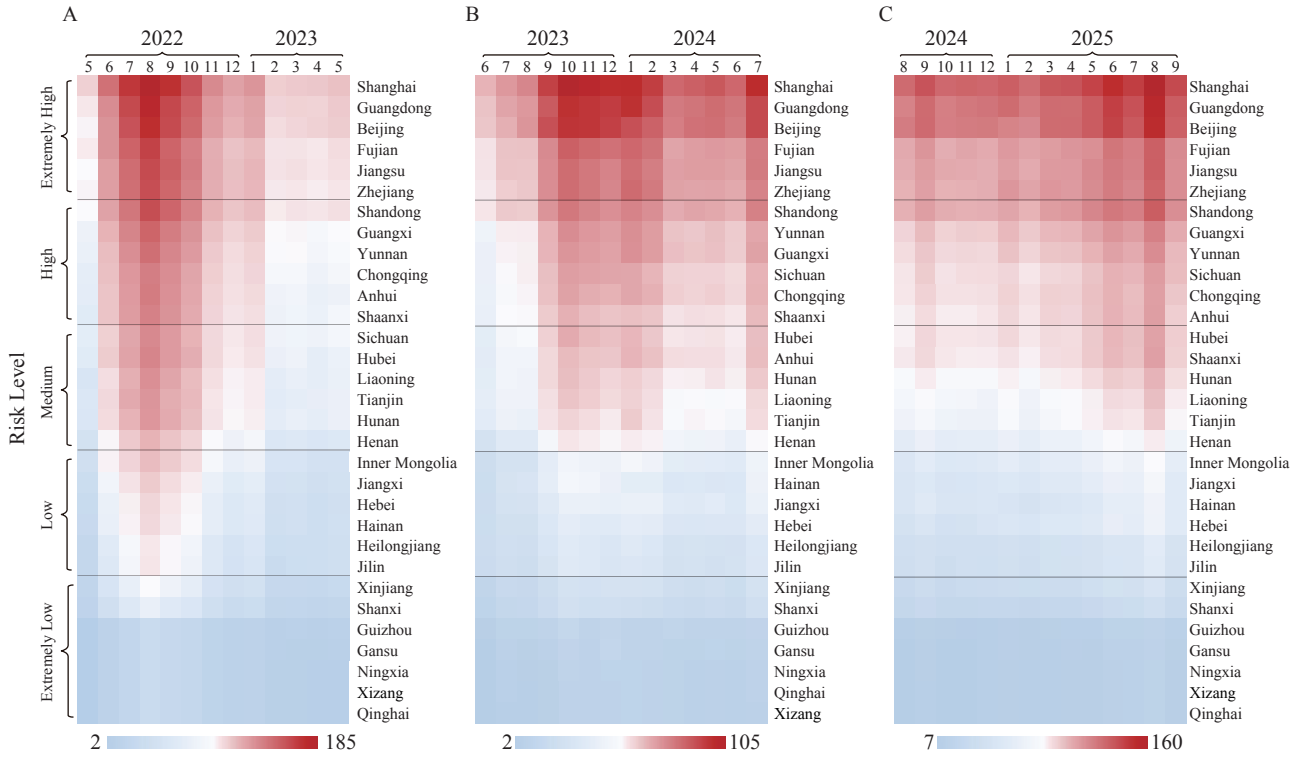
Source country to importing PLADs connectivity (C): C comprehensively considers aviation data and other factors potentially influencing connectivity, distinguishing between two aspects: the probability of importation from abroad and socioeconomic interactions. Use the following three parameters to calculate C by normalizing them to 0–1: air travel conditions (F), socio-geographic-mobility conditions (A), and history of similar disease imports from the source country (B).

A, B, and F are calculated as follows:

$$A, B, F = \sum_{i=1}^{n} a_i \times F_i \times b_i \quad (3)$$

A is calculated based on indicators F_{16}, F_{18}, F_{19}, F_{21} and F_{22}; B is computed using indicator F_{17}; F is calculated from indicators F_{15} and F_{20}. Where a_i denote the weight coefficient of the secondary indicators, b_i denote the weight coefficient of the tertiary indicators, and F_i represent the score of the tertiary indicators (on a scale from 1 to 5).

The F normalization formula is as follows:



SUPPLEMENTARY FIGURE S2. Importation risk of non-clade-specific across 31 PLADs in China.
Abbreviation: PLAD=provincial-level administrative division.

$$F' = \frac{F - F_{\min}}{F_{\max} - F_{\min}} \tag{4}$$

The same normalization method was used for A′ and B′.

The calculation of C is as follows:

$$C = A' + B' + F' \tag{5}$$

Details of the Framework Development

This system primarily evaluates the likelihood of infectious disease importation based on two key corrections: adjustments for the incidence rates in source countries and adjustments for the degree of connectivity between the source country and the PLAD in China.

Given the potential for underreporting of infectious disease cases in the source country, it is deemed necessary to correct the incidence rate using the incidence correction coefficient (S). This incidence rate correction coefficient S is determined by considering two factors: the epidemic detection capability of source country and the disease characteristics.

The epidemic detection capability of source country primarily considers the public health system of the source country, including factors such as the effective universal health coverage in source country (F_1), public health budget of source country (F_2), health workforce in source country (F_3), and pathogen diagnosis technology (F_4). In the process of establishing the evaluation system and creating the first-round questionnaire, we took into account the average education level of the source country to demonstrate the cognitive ability of the people in the source country with respect to diseases. Nevertheless, the questionnaire feedback indicated that two experts gave this indicator a weight of zero, suggesting that they regarded it as either unimportant or unable to reflect the intended aspect. Consequently, this indicator was deleted. Moreover, experts proposed adding the completeness of disease surveillance and early warning plans, systems, and capabilities. We hold that the Global Health Security Index can reflect this aspect to a certain degree. Consequently, the Global Health Security Index (F_5) has been newly incorporated. The gross domestic product (GDP) indicator of the source country was considered as potentially

indirectly reflecting the epidemic detection capability. However, in the second-round questionnaire, experts suggested deleting this indicator, as they believed it functions through public health budgets. We adopted this suggestion. Ultimately, the epidemic detection ability of the source country is represented by a total of five indicators, namely F_1 to F_5.

The characteristics of the disease considered include factors affecting detection and transmission, such as common symptoms (F_6), case fatality rate (F_7), herd susceptibility (F_8), incubation period (F_9), infectiousness during incubation period ($F10$), and communicable period (F_{11}). Additionally, the severity of the disease's impact may indirectly affect detection; hence, we also considered the classification of notifiable infectious disease (F_{12}) and the classification of emerging infectious disease (F_{13}). When formulating the evaluation framework and developing the initial questionnaire, the sequelae of the disease, whether it constitutes a Public Health Emergency of International Concern (PHEIC), and whether it is listed as a priority notifiable infectious disease were also considered in assessing the severity of the disease impact. However, each of these three indicators was assigned a weight of zero by two experts each, indicating that they deemed these indicators unimportant or not reflective. Consequently, these three indicators were removed. In the feedback from the first-round questionnaire, experts noted that the mode of transmission is an important characteristic of the disease and should be included as an indicator. Therefore, we added the main routes of disease transmission (F_{14}). Ultimately, the disease characteristics are represented by a total of nine indicators, namely F_6 to F_{14}.

The degree of connectivity between the source country and the PLAD in China determines the possibility of infectious disease import from another aspect. It consists of two parts: the possibility of importation from abroad and the socioeconomic interactions.

In terms of the possibility of importation from abroad, given that air travel serves as a significant mode of case importation, we considered aviation indicators such as the number of seats on flights from the source country to the local area (F_{15}) and the flight frequency from the source country to the local area (F_{20}). Additionally, we took into account other factors, including population migration index of the source country (F_{16}), history of similar disease importation from source country (F_{17}) and quantity of inbound overnight tourists received locally (F_{19}), to adjust for importation through non-aviation means. In establishing the assessment framework and developing the initial questionnaire, we also considered four additional indicators to correct for the likelihood of overseas importation: foreign direct investment from the source country, the number of personnel sent to contracted projects in the source country, the number of cases historically imported from the source country and the number of personnel dispatched for labor cooperation in the source country. We believed these indicators could potentially influence the risk of overseas importation. However, in the initial round of feedback from the questionnaire, each of these four indicators was assigned a weight of zero by at least two experts, indicating that they deemed the indicators unimportant or irrelevant. Consequently, we excluded these four indicators in the second round of questionnaire. Some experts suggested that the geographical location of the source country could influence the likelihood of overseas importation. Therefore, we added the geographical location of the source country (F_{18}) as a new indicator. Ultimately, the probability of importation from abroad is characterized by six indicators ranging from F_{15} to F_{20}.

The socioeconomic interactions component primarily aims to correct for importation through non-aviation means by assessing the degree of closeness between the local area and foreign regions. We considered the value of local imports and exports (F_{21}) and local international tourism revenue (F_{22}) as key indicators. In establishing the assessment framework and developing the initial questionnaire, we also considered the number of foreign businesses operating locally and the total foreign investment in the local area as indicators. However, four experts assigned a weight of zero to the number of foreign businesses, and three experts did the same for the total foreign investment, indicating that they deemed these indicators unimportant or irrelevant. Consequently, we excluded these two indicators in the second round of questionnaire. In the second round of feedback from the questionnaire, one expert noted that the value of local imports and exports may not fully align with population movements. While we found this suggestion to be valid, we decided to retain this indicator because no more suitable alternative was available, and it still provides insight into the local area's interactions with foreign regions. Consequently, the socio-economic exchanges are characterized by two indicators: F_{21} and F_{22}.

Methods and Applications

Automated Monkeypox Identification from Electronic Medical Records Using Large Language Models — Shenzhen City, Guangdong Province, China, 2023–2025

Diyang Xue[1,&]; Ye Ye[2,&]; Yu Wu[1]; Zhen Zhang[1]; Jianhua Lu[1]; Xuan Zou[1,#]; Qiuying Lv[3,#]

ABSTRACT

Introduction: Monkeypox (mpox) is an emerging infectious disease with diverse clinical manifestations that complicate early case identification. Traditional diagnostic methods relying on laboratory confirmation may delay detection during the early stages of infection. Clinical free-text data in electronic medical records (EMRs), such as chief complaints and histories of present illness, contain valuable information for early identification yet remain underutilized in routine surveillance. This study explored an automated approach to mpox identification from EMRs using large language models (LLMs).

Methods: We conducted a retrospective study of 239 individuals — 126 laboratory-confirmed mpox cases and 113 non-mpox controls — from multiple hospitals in Shenzhen, China. Free-text chief complaints and histories of present illness were processed using the DeepSeek-R1-14B large language model to extract structured clinical variables. Fourteen clinically relevant features were selected for model development. Naïve Bayes, logistic regression, and random forest classifiers were trained on these features, and a Jieba plus term frequency–inverse document frequency (TF-IDF) approach served as a baseline comparison. Qwen3-14B was also applied for feature extraction as a secondary comparison.

Results: mpox cases most frequently presented at fever clinics (55.6%), with considerable variation in care-seeking patterns. Models using DeepSeek-extracted features outperformed those based on traditional TF-IDF representations and Qwen3-14B, with logistic regression achieving the best performance [area under the receiver operating characteristic curve (AUROC), 0.927; accuracy, 87.5%]. DeepSeek demonstrated high accuracy (96.1%) in extracting relevant clinical symptoms, including fever, rash, pustules, and HIV infection.

Conclusion: LLM-based extraction of clinical features from routine text data represents a promising approach for automated mpox case identification and may support future intelligent surveillance and early warning systems.

Monkeypox (mpox) is a zoonotic disease caused by the monkeypox virus, a member of the *Orthopoxvirus* genus that also includes variola, vaccinia, and cowpox viruses (*1*). The disease presents with heterogeneous clinical manifestations, most commonly involving dermatologic and mucocutaneous lesions. Early identification of suspected cases is critical for timely isolation, epidemiologic investigation, and interruption of transmission.

In many clinical settings, mpox remains a relatively low-prevalence disease, and clinicians may have limited experience with its presentation. As a result, mpox may not be considered during the initial diagnostic evaluation, and patients may not be promptly referred for confirmatory testing. This gap in early clinical recognition is particularly concerning given the high transmissibility of mpox, as delayed or missed diagnoses can facilitate ongoing transmission and elevate public health risks.

Current diagnostic practices for mpox rely on epidemiological exposure assessment and clinical evaluation, followed by laboratory confirmation through nucleic acid amplification testing such as real-time or conventional polymerase chain reaction (PCR). Although these methods provide high diagnostic accuracy, laboratory confirmation typically occurs after the initial clinical encounter, limiting their utility for early case detection and real-time surveillance. These limitations underscore the need for complementary approaches that enable earlier identification of suspected mpox cases at the point of care.

Clinical narratives documented in electronic medical record (EMR) systems — including chief complaints,

physical examination findings, and histories of present illness — contain rich descriptive information that may facilitate earlier recognition of mpox. Ideally, EMR systems could automatically analyze these narratives and generate alerts for suspected cases.

As an initial investigation, we evaluated the feasibility of automatically identifying suspected mpox cases using unstructured clinical narratives. We developed machine learning models based on clinical features extracted by large language models (LLMs). Through contextual reasoning, LLMs can infer entity boundaries, semantic categories, and attributes — such as symptom presence or negation — from free-text clinical documentation, even without explicit rules. In contrast, traditional natural language processing (NLP) approaches, such as rule-based methods and bag-of-words models, including term frequency–inverse document frequency (TF-IDF) (*2*), require extensive manual curation, are sensitive to linguistic variability, and often fail to capture complex clinical semantics.

METHODS

Study Design and Data Collection

This retrospective, exploratory study used clinical records routinely collected from multiple hospitals in Shenzhen, China. The study included 239 individuals — 126 laboratory-confirmed mpox cases and 113 non-mpox controls — identified between July 18, 2023, and October 18, 2025. All mpox patients met the diagnostic criteria and had complete chief complaint and history of present illness documentation.

Controls were randomly selected from non-mpox patients who visited the same clinical department on the same day as each mpox case, ensuring comparable clinical context and healthcare-seeking patterns. Importantly, controls were not healthy individuals but rather patients presenting with various clinical complaints, including dermatologic conditions, infectious diseases, and other non-mpox diagnoses.

Because the study relied on free-text clinical narratives for feature extraction, records with missing or empty chief complaint and history of present illness fields could not be processed and were excluded. This exclusion resulted in a slight imbalance between cases and controls and may have introduced selection bias.

Given the emerging nature of mpox and the limited number of confirmed cases, this study was designed as an exploratory proof-of-concept investigation using all available eligible cases during the study period. A formal sample size calculation was not performed.

Large Language Model–based Feature Extraction

DeepSeek-R1-14B was deployed via an application programming interface (API) for prompt-based extraction of clinical information from free-text chief complaints and histories of present illness. A predefined prompt (Figure 1) instructed the model to perform entity recognition and attribute classification, identify symptoms, clinical signs, and relevant comorbidities, and assign binary presence or absence labels.

Feature extraction was conducted under deterministic inference by fixing the temperature parameter at 0 to ensure reproducibility and minimize stochastic variation. The model was applied in a zero-shot setting without task-specific fine-tuning. Raw model outputs were post-processed into standardized binary variables, with consistency checks applied to ensure conformity with the predefined schema. This process generated 98 candidate binary clinical variables per individual (Supplementary Material, available at https://weekly.chinacdc.cn/), spanning dermatologic manifestations, systemic symptoms, gastrointestinal symptoms, and relevant medical history.

Feature Selection Based on Clinical Knowledge

From the 98 candidate variables extracted by the LLM, feature selection was guided by existing clinical guidelines and expert consultation with infectious disease and public health specialists (*3*–*4*). Given the exploratory nature of the study, the relatively small sample size, and the need for model interpretability, features were selected manually rather than through automated methods. Additional analyses using all 98 variables and least absolute shrinkage and selection operator (LASSO)-based feature selection were conducted to evaluate the impact of manual feature selection.

Selection prioritized variables associated with characteristic mpox lesions, high-risk populations, and sufficient prevalence in the study population. Fourteen variables were retained for model development: inguinal rash, inguinal swelling and pain, perianal papules, perianal pustules, perianal pain, fever, rash, papules, pustules, AIDS, hepatitis, diarrhea, dull lower abdominal pain, and skin infections.

```
("任务：从中文病历片段中判断腹股沟皮疹、腹股沟肿痛、肛周丘疹、肛周脓疱、肛周疼痛、发热、皮疹、丘疹、脓疱、艾滋病、肝炎、腹泻、下腹部隐痛、皮肤感染及其他症状。\n"
"输出规则：\n"
"1)每个症状一条记录，text present (是否)\n"
"2)若 present = 是，再给 acute(是否)。出现急性\慢性等词据此判断，未提及可留空。\n"
"3)可选字段 duration，如近两天等。\n"
"4)仅输出 Json，必须符合 Json scheme。\n"
"示例\n"
"输入：患者发热三天伴皮疹。\n"
"期望:{\" symptoms\":["
"{"text\:发热, \"present\:\"是\", \" acute\:\"是\", \" duration \": \"三天\"} "
"{"text\:皮疹, \"present\:\"是\", \" acute\:\"是\", \" duration \": \"三天\"} "
"]}\n"
f "现在输入:{text}"
)
```

```
("Task: From Chinese medical record fragments, determine inguinal rash, inguinal swelling and pain, perianal papules, perianal pustules, perianal pain, fever, rash, papules, pustules, AIDS, hepatitis, diarrhea, dull lower abdominal pain, skin infections, and other symptoms.\n"
"Output rules:\n"
"1) One record per symptom, text present (yes/no) \n"
"2) If present = yes, then provide acute (yes/no). Determine based on the appearance of terms such as acute\chronic; if not mentioned, it may be left blank.\n"
"3) Optional field duration, such as past two days, etc. \n"
"4) Output Json only, must conform to Json schema. \n"
"Example\n"
"Input: The patient has had fever for three days accompanied by rash.\n"
"Expected: {\" symptoms\":["
"{"text\:fever, \"present\:\"yes\", \" acute\:\"yes\", \" duration \": \"three days\"} "
"{"text\:rash, \"present\:\"yes\", \" acute\:\"yes\", \" duration \": \"three days\"} "
"]}\n"
f "Current input: {text}"
)
```

FIGURE 1. Prompt for symptom extraction from Chinese clinical text for monkeypox cases and controls.
Note: left panel: original Chinese; right panel: English translation.

Model Construction and Baseline Comparison

The dataset was randomly split into training and validation sets at an 8:2 ratio. Using the same 14 LLM-extracted variables, three classification models were developed: naïve Bayes, logistic regression, and random forest. No additional internal feature selection was applied, ensuring comparability across models.

For the naïve Bayes model, binary variables were modeled under a Bernoulli distribution assumption. Logistic regression was implemented with L2 regularization, and random forest models consisted of ensembles of decision trees.

To further assess model robustness and reduce the risk of overfitting, five-fold cross-validation was additionally performed on the training dataset for the primary model (logistic regression using DeepSeek-extracted features).

To evaluate the added value of LLM-based feature extraction, a traditional NLP baseline was implemented using the Jieba word segmentation tool (*5*) combined with TF-IDF feature representation. Naïve Bayes, logistic regression, and random forest models were constructed using the same workflow for comparison.

In addition to DeepSeek-R1-14B, the Qwen3-14B large language model (*6*) was applied using the same prompt, inference settings, and feature selection procedures to provide a comparative assessment of different LLM-based feature extraction approaches. Model performance was evaluated using area under the receiver operating characteristic curve (AUROC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Brier score for calibration assessment.

RESULTS

Clinical Care-Seeking Patterns of Monkeypox Cases

Mpox cases presented across multiple clinical departments, most commonly fever clinics (55.6%), followed by infectious disease outpatient clinics (16.7%) and HIV/AIDS clinics (9.5%) (Table 1). Dermatology-related clinics accounted for 11.1% of initial visits, with the remaining cases distributed across emergency departments and other outpatient services.

This distribution reflects heterogeneous care-seeking

TABLE 1. Distribution of mpox cases across clinical departments.

Clinical Department	Number of Cases (*n*)	Proportion (%)
Fever clinic	70	55.6
Infectious disease outpatient clinic	21	16.7
HIV/AIDS clinic	12	9.5
Dermatology outpatient clinic	10	7.9
Research follow-up clinic	6	4.8
Dermatology and venereology clinic	3	2.4
Emergency department outpatient clinic	1	0.8
Dermatologic aesthetic and laser clinic	1	0.8
Wound, ostomy, and continence nursing specialty clinic	1	0.8
Proctology night clinic	1	0.8
Total	126	100.0

Abbreviation: mpox=monkeypox; HIV=human immunodeficiency virus; AIDS=acquired immunodeficiency syndrome.

patterns and highlights the limitations of department-specific surveillance strategies for early mpox case identification.

Accuracy of DeepSeek LLM in Clinical Symptom Extraction

As shown in Table 2, DeepSeek-R1-14B demonstrated strong performance in extracting 14 mpox-related clinical symptoms from unstructured clinical text, achieving a precision of 0.771, recall of 0.919, F1 score of 0.839, and accuracy of 0.961. The model showed particularly high precision and recall for commonly documented symptoms, including fever, rash, pustules, and HIV infection. Overall, these results suggest that DeepSeek can reliably extract clinically relevant symptom information from routine clinical narratives.

Comparative Classification Performance Across Feature Extraction Methods

As shown in Table 3, models using LLM-extracted features consistently outperformed those based on traditional Jieba plus TF-IDF representations across multiple evaluation metrics, including AUROC, accuracy, sensitivity, specificity, PPV, NPV, and calibration (Brier score).

Among all models, logistic regression using DeepSeek-R1-14B–extracted features achieved the best overall performance, with the highest AUROC (0.927), accuracy (87.5%), balanced sensitivity (88.0%) and specificity (87.0%), and the lowest Brier score (0.106), indicating superior calibration.

Models based on Qwen3-14B–extracted features demonstrated moderate performance, generally achieving higher sensitivity but lower specificity and less favorable calibration compared with DeepSeek-based models. In contrast, Jieba plus TF-IDF–based models showed the weakest overall performance, with reduced specificity and PPV.

A five-fold cross-validation on the training dataset yielded a mean AUROC of 0.904, consistent with the performance observed on the independent validation set.

Feature Selection Analysis

To evaluate the impact of manual feature selection, additional analyses were conducted using all 98 LLM-extracted variables and a LASSO logistic regression for data-driven feature selection. Logistic regression using all 98 variables achieved an AUROC of 0.910 and an accuracy of 85.4%, while the LASSO-based model achieved an AUROC of 0.902 and an accuracy of 87.5%. These results were comparable to the 14-feature clinically selected model and did not demonstrate clear improvement.

LASSO-based feature selection identified several clinically relevant dermatologic and systemic symptoms associated with mpox, suggesting that meaningful signals can be captured directly from high-dimensional feature space. However, models based on clinically selected features demonstrated more stable and interpretable performance, particularly given the relatively small sample size.

Error Case Analysis

Error case analysis revealed several patterns underlying model misclassification. False-positive cases were commonly associated with nonspecific

TABLE 2. Performance evaluation of DeepSeek-R1-14B for extracting mpox-related clinical symptoms.

Symptom	TP	FP	TN	FN	Precision	Recall	F1 Score	Accuracy
Inguinal rash	10	0	228	0	1.000	1.000	1.000	1.000
Inguinal swelling and pain	2	0	236	0	1.000	1.000	1.000	1.000
Perianal papules	9	7	222	0	0.563	1.000	0.720	0.970
Perianal pustules	7	2	227	2	0.778	0.778	0.778	0.983
Perianal pain	6	1	231	0	0.857	1.000	0.923	0.996
Fever	86	11	141	0	0.887	1.000	0.940	0.954
Rash	97	12	128	1	0.890	0.990	0.940	0.945
Papules	31	63	143	1	0.330	0.969	0.492	0.731
Pustules	34	3	182	19	0.919	0.624	0.756	0.908
AIDS	37	1	199	1	0.974	0.974	0.974	0.992
Hepatitis	7	1	229	1	0.875	0.875	0.875	0.992
Diarrhea	10	0	225	3	1.000	0.769	0.870	0.987
Dull lower abdominal pain	1	0	237	0	1.000	1.000	1.000	1.000
Skin infections	3	0	233	2	1.000	0.600	0.750	0.992
Total	340	101	2,861	30	0.771	0.919	0.839	0.961

Note: Precision=TP / (TP + FP); Recall=TP / (TP + FN); F1=2 × Precision × Recall / (Precision + Recall); Accuracy=(TP + TN) / (TP + FP + TN + FN).
Abbreviation: mpox=monkeypox; AIDS=acquired immunodeficiency syndrome; TP=true positive; FP=false positive; TN=true negative; FN=false negative.

TABLE 3. Comparison of classification performance using DeepSeek-R1-14B, Qwen3-14B, and Jieba + TF-IDF features.

Feature extraction	Model	AUROC	Accuracy, %	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Brier score
DeepSeek-R1-14B	Naive Bayes	0.910	85.4	88.0	82.6	84.6	86.4	0.141
DeepSeek-R1-14B	Logistic Regression	0.927	87.5	88.0	87.0	88.0	87.0	0.106
DeepSeek-R1-14B	Random Forest	0.890	85.4	88.0	82.6	84.6	86.4	0.118
Qwen3-14B	Naive Bayes	0.770	72.9	76.0	69.6	73.1	72.7	0.271
Qwen3-14B	Logistic Regression	0.838	81.2	92.0	69.6	76.7	88.9	0.147
Qwen3-14B	Random Forest	0.861	81.2	92.0	69.6	76.7	88.9	0.141
Jieba + TF-IDF	Naive Bayes	0.790	68.8	80.0	56.5	66.7	72.2	0.189
Jieba + TF-IDF	Logistic Regression	0.819	75.0	80.0	69.6	74.1	76.2	0.188
Jieba + TF-IDF	Random Forest	0.877	79.2	92.0	65.2	74.2	88.2	0.152

Abbreviation: AUROC=area under the receiver operating characteristic curve; PPV=positive predictive value; NPV=negative predictive value; TF-IDF=term frequency–inverse document frequency.

dermatologic symptoms, such as erythema or papules, that were over-interpreted as characteristic mpox lesions in the absence of systemic indicators. For example, a patient presenting with perianal erythema and itching was clinically diagnosed with localized eczema, yet the model assigned a high probability of mpox based on the anatomical location of the symptoms.

In contrast, false-negative cases primarily involved patients presenting with systemic symptoms such as fever and myalgia but lacking documented characteristic rashes or pustules at the initial encounter. These findings suggest that further refinement is needed to improve differentiation between mpox and common dermatologic conditions with overlapping presentations.

Comparison Between DeepSeek and Qwen3 Feature Extraction

To further investigate the performance differences between DeepSeek-R1-14B and Qwen3-14B, we conducted a qualitative comparison of their feature extraction outputs using identical prompts and inference settings.

Although both models were capable of extracting structured clinical variables from free-text records,

DeepSeek demonstrated greater robustness in handling ambiguous or indirectly expressed symptoms. For example, when clinical narratives described "perianal discomfort" or "localized irritation," DeepSeek more consistently inferred dermatologic involvement, whereas Qwen3 frequently failed to capture these signals when explicit terms such as "rash" or "pustules" were absent.

Additionally, Qwen3 exhibited greater variability in feature extraction consistency across similar clinical descriptions, occasionally producing incomplete or less reliable binary representations. These findings suggest that stronger contextual understanding and instruction-following capability may contribute to the superior downstream classification performance observed with DeepSeek.

DISCUSSION

Main Findings

This study demonstrated that LLM-extracted clinical features can improve mpox case identification from routine clinical text, with DeepSeek-R1-14B yielding the strongest performance among the evaluated models. The logistic regression model built on DeepSeek-derived features achieved an AUROC of 0.927 and an accuracy of 87.5%, indicating that the model can capture clinically relevant signals from unstructured medical narratives.

Beyond overall model performance, the favorable results may also reflect the clinical relevance of the symptom features used for classification. The selected features represent common and emerging mpox presentations observed in recent clinical practice, particularly dermatologic manifestations involving the inguinal and perianal regions, which are frequently described in routine medical records (*7*). Local symptoms such as inguinal swelling and perianal pain capture inflammatory responses — including lymphadenopathy and rectal involvement — that may occur even when typical skin lesions are limited or absent. Systemic and gastrointestinal symptoms, including fever and diarrhea, further account for heterogeneous clinical presentations and early-stage disease. Additionally, hepatitis reflects systemic involvement or coexisting conditions observed in some patients, particularly in the context of immune dysfunction or concurrent infections (*8*). Moreover, incorporating immune status–related information such as HIV/AIDS enables the model to capture clinically meaningful variation in symptom burden and disease severity (*9*). Together, these features provide a clinically grounded representation of mpox manifestations and explain the model's ability to extract discriminative signals from unstructured clinical text.

Clinical Translation

The observed performance of DeepSeek highlights the potential utility of LLMs in supporting infectious disease surveillance and clinical decision-making. Because the proposed framework relies solely on routinely collected EMR data — including chief complaints and histories of present illness — it may be feasible to integrate into existing hospital information systems without requiring additional data collection or workflow disruption.

Automated screening based on clinical narratives may facilitate earlier identification of suspected mpox cases before laboratory confirmation, particularly in settings where patients present across multiple clinical departments. In addition, the LLM-based approach demonstrated some ability to capture semantically related symptom descriptions beyond explicit keywords, suggesting potential advantages in handling heterogeneous clinical documentation.

Ethical and privacy considerations are also important for practical implementation. This study used retrospective, de-identified clinical data, and any future deployment would require strict adherence to data protection regulations and institutional governance. Importantly, the proposed system is intended to support — rather than replace — clinical decision-making, and human oversight remains essential.

Finally, this study was based on retrospective data and did not include temporal stratified analysis or prospective validation. Future work should focus on evaluating model performance across different time periods and conducting prospective validation in real-world clinical settings to assess generalizability and robustness.

Limitations

Several limitations should be considered when interpreting these findings. First, this study should be regarded as a proof-of-concept investigation based on a relatively small retrospective dataset, which may limit generalizability and increase the risk of overfitting despite consistent five-fold cross-validation performance. External validation across different healthcare systems and prospective evaluation in real-world settings were not performed.

Second, the approach relies heavily on free-text clinical documentation, which may be incomplete, nonstandardized, or ambiguous. Although DeepSeek demonstrated some capability in handling indirectly expressed symptoms, performance may decrease when clinical descriptions are highly vague or atypical. In addition, the exclusion of records with incomplete narratives may have introduced selection bias.

Third, clinically guided manual feature selection may introduce verification bias by constraining the model to predefined symptom patterns. Although additional analyses using all 98 variables and LASSO-based feature selection demonstrated comparable performance, clinically selected features provided more stable and interpretable results.

Finally, the baseline comparison in this study was limited to Jieba plus TF-IDF feature extraction. More advanced traditional clinical NLP approaches, such as rule-based named entity recognition systems or neural sequence labeling models, may provide stronger baselines for future evaluation.

Future Directions

Despite these limitations, the proposed framework demonstrates the feasibility of using LLMs for automated extraction of clinically relevant information from routine medical narratives. Because the approach operates in a zero-shot manner without task-specific annotation or model fine-tuning, it may offer practical advantages for rapid deployment in emerging infectious disease surveillance settings.

Future studies should evaluate the framework using larger multicenter datasets, external validation cohorts, and prospective real-world implementation. Incorporating additional data modalities — such as medical imaging or population-level surveillance data — may further improve early identification of atypical presentations. More broadly, similar approaches may be adaptable to other emerging or reemerging infectious diseases that rely heavily on narrative clinical documentation.

The DeepSeek-R1-14B large language model, combined with clinically informed feature selection, demonstrated strong performance in automated mpox case identification using routine clinical text. These findings suggest that LLM-based processing of clinical narratives represents a promising and scalable approach for supporting infectious disease surveillance. Despite the limitations of this proof-of-concept study, the proposed framework may be adaptable to other infectious diseases and integrated into existing surveillance systems following further validation.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the Shenzhen Medical Research Fund (Grant No. E24010011, E24010010) and the Shenzhen Science and Technology Program (Grant No. SYSPG20241211173921049). The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

doi: 10.46234/ccdcw2026.102

[#] Corresponding authors: Qiuying Lv, lvqiuying@wjw.sz.gov.cn; Xuan Zou, zouxuan@wjw.sz.gov.cn.

[1] Shenzhen Center for Disease Control and Prevention, Shenzhen City, Guangdong Province, China; [2] School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen City, Guangdong Province, China; [3] School of Public Health, Southern Medical University, Guangzhou City, Guangdong Province, China.
[&] Joint first authors.

Submitted: January 23, 2026
Accepted: May 08, 2026
Issued: May 15, 2026

REFERENCES

1. Mitjà O, Ogoina D, Titanji BK, Galvan C, Muyembe JJ, Marks M, et al. Monkeypox. Lancet 2023;401(10370):60 – 74. https://doi.org/10.1016/S0140-6736(22)02075-X.
2. Salton G, Buckley C. Term-weighting approaches in automatic text retrieval. Inf Process Manage 1988;24(5):513 – 23. https://doi.org/10.1016/0306-4573(88)90021-0.
3. National Health Commission of the People's Republic of China. Guidelines for the diagnosis and treatment of monkeypox (2022 version). Chin J Viral Dis 2022;12(4):255 – 6. https://doi.org/10.16505/j.2095-0136.2022.0040.
4. World Health Organization. Clinical management and infection prevention and control for mpox: living guideline, May 2025. Geneva: World Health Organization; 2025. http://dx.doi.org/10.2471/B09434.
5. Sun J. Jieba: Chinese text segmentation. 2012. https://github.com/fxsjy/jieba. [2025-11-15].
6. Yang A, Li AF, Yang BS, Zhang BC, Hui BY, Zheng B, et al. Qwen3 technical report. 2025. https://arxiv.org/abs/2505.09388. [2025-12-03].
7. Pittman PR, Martin JW, Kingebeni PM, Tamfum JJM, Wan QW, Reynolds MG, et al. Clinical characterization of human monkeypox infections in the Democratic Republic of the Congo. medRxiv 2022. http://dx.doi.org/10.1101/2022.05.26.22273379. [2025-12-03].
8. Calin R, Périllaud-Dubois C, Marot S, Kerrou K, Peytavin G, Bachir M, et al. Mpox hepatic and pulmonary lesions in HIV/hepatitis B virus co-infected patient, France. Emerg Infect Dis 2024;30(11):2445 – 7. https://doi.org/10.3201/eid3011.241331.
9. Saldana CS, Kelley CF, Aldred BM, Cantos VD. Mpox and HIV: a narrative review. Curr HIV/AIDS Rep 2023;20(4):261 – 9. https://doi.org/10.1007/s11904-023-00661-1.

SUPPLEMENTARY MATERIAL

A total of 98 clinical features were extracted from electronic medical records using DeepSeek-R1-14B. The 98 clinical features extracted were: inguinal rash, inguinal swelling and pain, perianal papules, perianal pustules, perianal pain, fever, rash, papules, pustules, acquired immunodeficiency syndrome (AIDS), hepatitis, diarrhea, dull lower abdominal pain, skin infections, dizziness, fatigue, hyperlipidemia, perianal herpes, cough, sore throat, nasal congestion, vomiting, abdominal pain, dysuria, perianal rash, perianal abscess, syphilis, headache, myalgia, chills, rigors, human immunodeficiency virus positive (HIV-positive), diabetes mellitus, genital ulcer, muscle and joint pain, rhinorrhea, cough with tinnitus, soft tissue infection, penile swelling and pain, pruritus, pain, productive sputum, itching sensation, perianal inflammation, herpesvirus infection of the genital and genitourinary tract, herpes, generalized muscle and joint pain, cough with white sputum, facial swelling, difficulty in defecation, mild myalgia, recurrent condyloma acuminatum, painful rash, bronchitis, respiratory tract infection, cough-variant asthma, renal insufficiency, nausea and vomiting, generalized muscle and joint discomfort, dry throat, dry cough with scant sputum, gingival pain, nasal congestion and rhinorrhea, pulmonary tuberculosis, night sweats, dyspnea, chest pain, hemoptysis, parasitic infection, cough with yellow sputum, neurosyphilis, poor appetite, poor general condition, chest mass, multiple lipomas, poor mental status and appetite, poor sleep, parasitic disease, cough with yellow mucoid sputum, venous thrombosis, Grade 3 hypertension (very high risk), abdominal pain and diarrhea, bacterial infection, viral warts, human papillomavirus (HPV) infection, abdominal distension, bilateral breast tenderness, mild headache, chest tightness, productive cough, nausea, throat itching, rhinorrhea and nasal congestion, chronic hepatitis B, poor postoperative wound healing, androgenetic alopecia, and kidney stones.

Youth Editorial Board

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).

CHINA CDC WEEKLY

CCDC weekly

Vol. 8 No. 20 May 15, 2026

中国疾病预防控制中心周报(英文)

Responsible Authority
National Disease Control and Prevention Administration

Sponsor
Chinese Center for Disease Control and Prevention &
Chinese Academy of Preventive Medicine

Editor-in-Chief
Jianwei Wang

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

Printing: Beijing Kexin Printing Co., Ltd

Complimentary Access

CSSN
ISSN 2096-7071 (Print)
ISSN 2097-3101 (Online)
CN 10-1629/R1