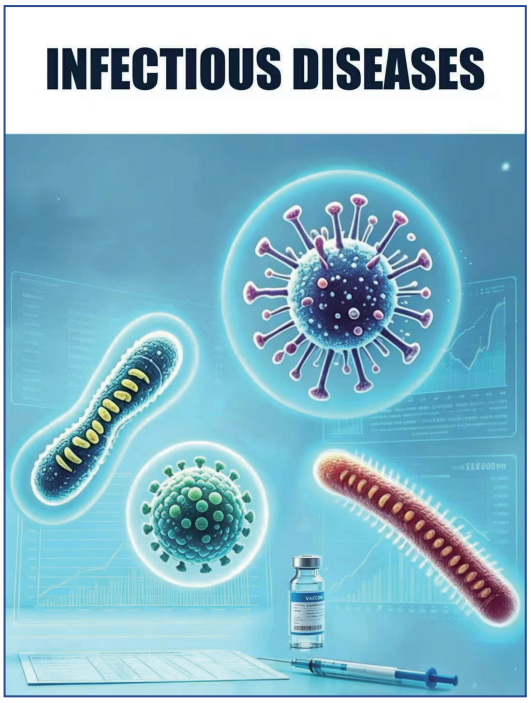
ISSN 2096-7071 (Print)
ISSN 2097-3101 (Online)
CN 10-1629/R1

CHINA CDC WEEKLY

Vol. 8 No. 16 Apr. 17, 2026

中国疾病预防控制中心周报(英文)



ISSN 2096-7071

China CDC Weekly

Editorial Board

Editorial Office

Cover Image: credited from the designer Jiaying Wang.

Vital Surveillances

Surveillance and Analysis of Plague Epidemic — China, 2010–2024

Qi Bing[1,&]; Shaoxian Zhang[1,&]; Zhao Wang[1]; Cheng Ju[1]; Shiwen Zhang[1]; Na Ma[1]; Bin Zhao[1]; Xiaoheng Yao[1]; Cheng Xu[1]; Kuidong Shao[1,#]

ABSTRACT

Introduction: Plague, caused by *Yersinia pestis*, is a class A infectious disease in China. In order to address the practical requirements of plague prevention and control, this study conducted a systematic analysis based on the national surveillance data from 2010 to 2024.

Methods: Data on human and animal plague epidemics were collected from the National Plague Prevention and Control Management Information System and analyzed using descriptive epidemiological methods.

Results: As of 2024, 12 types of natural plague foci have been identified in China. Between 2010 and 2024, 33 human plague cases and 14 deaths were reported nationwide. Geographically, cases were concentrated in Inner Mongolia (15, 45.5%), Xizang (8, 24.2%), and Gansu (6, 18.2%). By focus type, cases originated primarily from Marmota himalayana (16, 48.5%), *Meriones unguiculatus* (15, 45.5%), and *Rattus flavipectus* (2, 6.1%). Pneumonic plague was most common (15, 45.5%), followed by bubonic (10, 30.3%), septicemic (7, 21.2%), and intestinal plague (1, 3.0%). Animal plague epidemics persisted nationwide, with active outbreaks in foci of *M. himalayana, Marmota baibacina, Spermophilus undulatus*, and *Meriones unguiculatus*. Bacterial isolation positivity rates in animals and insects were 0.05%–0.13% and 0.04%–0.11%, respectively, while Indirect Hemagglutination Assay (IHA) and Reverse Indirect Hemagglutination Assay (RIHA) positivity rates were 0.08%–0.25% and 0.09%–2.18%, respectively. Currently, Polymerase Chain Reaction (PCR) testing has been implemented in some regions, with positivity rates of 0.80% in animals and 1.59% in insects in 2024.

Conclusions: The current plague epidemic situation in China is characterized by "sporadic human cases, active animal epidemics in some foci." It is recommended to strengthen targeted surveillance, promote PCR-based screening, and enhance cross-provincial coordination.

Plague is a zoonotic disease caused by *Yersinia pestis* (*1*–*2*). Its natural cycle relies primarily on rodents and fleas, and it can infect humans under specific conditions (*1*,*3*). China is one of the countries with the most diverse types and widespread distribution of natural plague foci, encompassing various ecological zones such as alpine grasslands, desert steppes, mountainous areas, and agro-pastoral ecotones (*4*). Since the 1980s, China has implemented a multilevel nationwide surveillance system and maintained a high level of preparedness for rapid epidemic responses (*5*). However, recent factors, such as ecological restoration, climate change, shifting human activity, evolving epidemic trends, and changing ecological characteristics of foci, have introduced new dynamics. These changes pose greater challenges to existing prevention and control systems (*6*). By integrating national surveillance data from 2010 to 2024, this study provides a comprehensive characterization of 15 years of plague epidemiology, focus evolution, and diagnostic technology use in China, with the goal of informing evidence-based control strategies.

METHODS

This study used surveillance data from the National Plague Prevention and Control Management Information System for 2010–2024. Descriptive epidemiological methods analyzed the dynamics of natural foci and human and animal epidemics in China. All surveillance activities, case definitions, laboratory methods, and data reporting strictly followed the National Plague Surveillance Program (2024 Revision) (*5*). Focus data covered spatial distribution and area changes in affected counties. Surveillance included fixed and mobile sites across 24 provincial-level administrative divisions (PLADs) and the Xinjiang Production and Construction Corps (XPCC), covering various ecological types such as alpine grasslands, desert steppes, mountainous areas,

and agro-pastoral ecotones. Human case data included clinical classification, outcomes, exposure history, and epidemiological information. Case definition and classification followed the *Diagnostic Criteria for Plague (WS 279-2008)* (*7*). Animal epidemic data included pathogenic, immunological, and molecular test results for host animals and vectors. To evaluate surveillance effectiveness, annual positive rates (positive sample numbers / total tested ×100%) was calculated for various testing methods. Data organization and statistical analyses were performed using Microsoft Excel 2019 (Microsoft Corporation, Redmond, Washington, USA).

RESULTS

Distribution of Natural Plague Foci

China has currently identified 12 types of natural foci. Between 2010 and 2024, 19 counties with new plague epidemic foci were identified. These are primarily located in Xizang, Sichuan, Xinjiang, Inner Mongolia and Yunnan PLADs. The *Marmota himalayana* focus (predominantly alpine habitat) showed the highest increase, with 13 new counties (13/19, 68.4%; 9 in Xizang, 4 in Sichuan). The *Rhombomys opimus* plague focus added two counties (10.5%), while the *Meriones unguiculatus*, *Marmota baibacina-Spermophilus undulatus, Marmota caudata*, and the *Apodemus chevrieri-Eothenomys miletus* foci added one (5.3%). These findings indicated that spatial expansion and belt-like intensification of natural plague foci in China are primarily concentrated in the high-altitude Xizang Plateau.

Characteristics of Human Plague

Between 2010 and 2024, 25 indigenous human plague outbreaks were reported nationwide, involving 33 cases, with 19 recoveries and 14 deaths (Figure 1). Geographically, cases were concentrated in Inner Mongolia (15, 45.5%), Xizang (8, 24.2%), and Gansu (6, 18.2%; Figure 2) PLADs. By focus type, cases originated primarily from the *Marmota himalayana* (16, 48.5%), *Mongolian gerbil* (15, 45.5%), and Asian house rat foci (2, 6.1%). Pneumonic plague was predominant (including secondary plague; 15, 45.5%), followed by bubonic (10, 30.3%), septicemic (including secondary septicemic plague; 7, 21.2%), and intestinal plague (1, 3.0%). Pneumonic plague was responsible for intrahousehold transmission and clustered outbreaks, as observed in Lang County (2010) and Sonid Right Banner (2023). Cases were predominantly male (24, 72.7%), with 9 females (27.3%). Age ranged from 3 years to 70 years, concentrated in the 20–59 age group (28, 84.8%), primarily young and middle-aged laborers. Occupational distribution was highly concentrated among herders and related occupations (27, 81.8%). The routes of infection included direct skinning or contact with infected animals, occupational exposure within plague foci, household transmission, and imported cases or cross-regional healthcare-seeking.

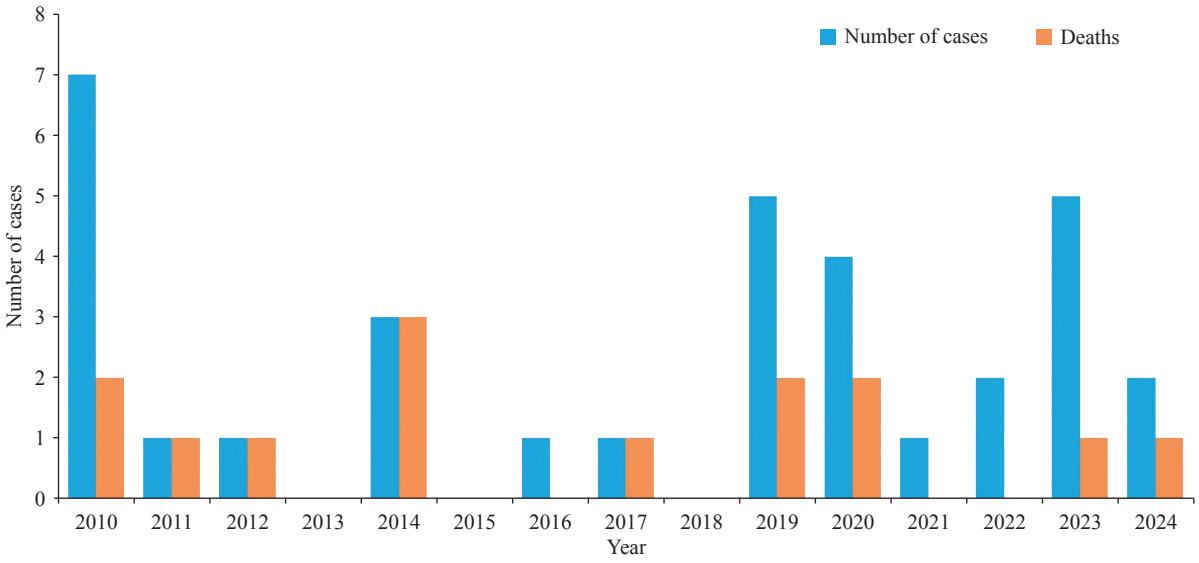


FIGURE 1. Annual number of reported human plague cases and deaths in China, 2010–2024.

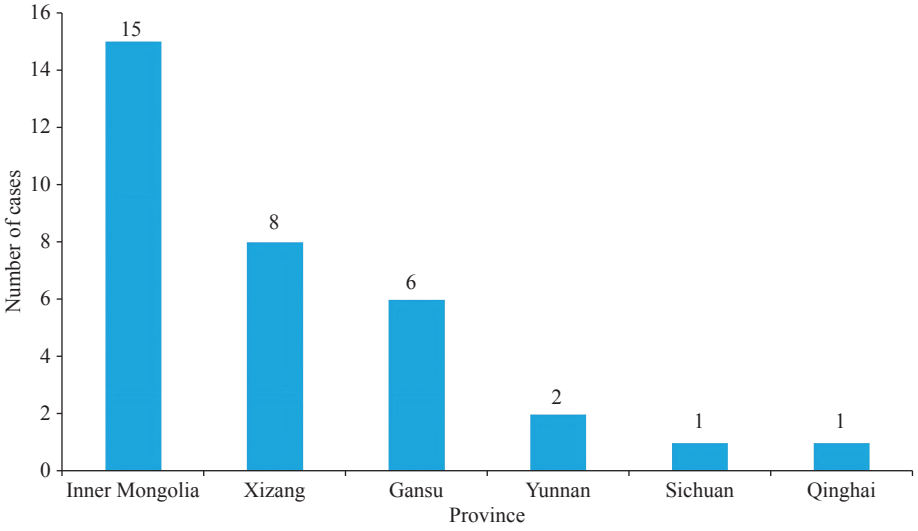


FIGURE 2. Regional distribution of human plague cases in China, 2010–2024.

Characteristics of Animal Plague

Animal plague epidemics persisted nationwide from 2010 to 2024, with significant variations in intensity across different foci. In the *M. himalayana focus*, pathogen isolation positivity rates were approximately 1.07% in host animals (957 positive strains) and 0.54% in vector insects (161 positive pools). In *Marmota baibacina* and *Spermophilus undulatus*, rates were approximately 0.20% in hosts (107 positive strains) and 0.16% in vectors (89 positive pools), indicating comparable sustained activity. In the *Meriones unguiculatus* focus, rates were approximately 0.16% in hosts (377 positive strains) and 0.19% in vectors (137 positive pools). A notable annual peak occurred in 2019, with a host animal positivity rate of approximately 0.52% (96 positive strains). Since then, animal epidemics have shown a developmental trend from east to west and north to south, with human cases reported annually for 6 consecutive years. Animal plaque has been continuously reported since the discovery of the *Microtus fuscus* focus. The *Rhombomys opimus* focus in the Junggar Basin exhibited a consecutive annual outbreak pattern, while sporadic epidemics occurred in the *Apodemus chevrieri-Eothenomys miletus* focus. In contrast, foci involving the *Spermophilus dauricus* and *Marmota sibirica* were quiescent. However, even in quiescent foci, sporadic positive detection underscores the need for continuous active surveillance.

Methodological analysis showed that pathogen isolation positive rates remained low over the long term. Host isolation positive rates ranged from 0.05% (2023) to 0.13% (2010), and vector rates from 0.04% (2012) to 0.11% (2010). For immunological detection, the indirect hemagglutination assay (IHA) positivity rate ranged from 0.08% (2016–2022) to 0.25% (2013). The reverse indirect hemagglutination assay (RIHA) positivity rate remained mostly below 1%, although it reached 2.18% in 2013. In 2024, animal and insect PCR-positive rates were 0.80% and 1.59%, respectively, both significantly higher than pathogen isolation results during the same period (Figure 3).

DISCUSSION

Analysis of Factors for Natural Plague Foci Expansion

Ecological Factors: The northward shift of isohyets (*8*) and the "warming and wetting" trend (*9–11*) have expanded the ecological space for host animals, driving the expansion of plague foci along climatic-ecological transition zones, a mechanism supported by the spatial clustering of newly affected counties.

Enhanced Surveillance Capability: The systematic strengthening of surveillance network technology and coverage between 2010 and 2024 has significantly improved the sensitivity of identifying potential risk areas.

Analysis of Human Plague Epidemiological Characteristics

Surveillance data from 2010 to 2024 indicated that

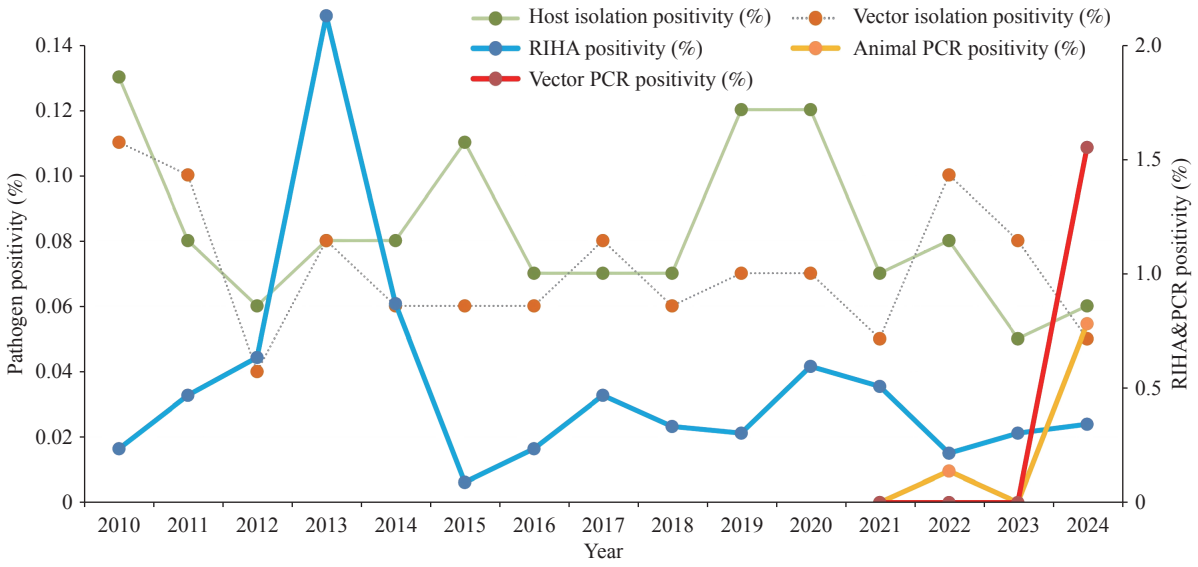


FIGURE 3. Surveillance positivity rates for major host animals and vector insects in China, 2010–2024.
Abbreviation: PCR=polymerase chain reaction; RIHA=reverse indirect hemagglutination assay.

human plague in China remained at a low level of sporadic occurrence. Cases were mostly sporadic and primarily involved "first-generation cases" resulting from direct contact (hunting, skinning, eating, and biting). Existing prevention and control systems have proven effective in blocking human-to-human transmission.

Geographically, human cases were concentrated in Inner Mongolia, Xizang, and Gansu (>80% of cases). By focus type, cases were predominantly distributed in the *Marmota himalayana* (48.5%) and *Meriones unguiculatus* (45.5%) foci, confirming the overlap between high-risk geographic areas and specific host foci. Furthermore, outbreaks in new areas resulting from dynamic changes in plague foci, such as the 2012 Litang County incident in Sichuan, highlight the potential risks posed by marginal zones and climatic anomalies (*12*–*13*).

Between 2010 and 2024, multiple cross-regional transmission events occurred in China, typified by the 2019 imported case from Inner Mongolia to Beijing (*14*). This underscores the potential risk of plague spreading from traditionally endemic areas to densely populated metropolitan regions. Therefore, the comprehensive measures encapsulated in the "Five Earlies and One Localization" principle, early detection, early reporting, early diagnosis, early isolation, early treatment, and localization of treatment, must be strictly implemented to rapidly interrupt transmission chains and control outbreaks.

Animal Plague and Surveillance Technologies

Surveillance data from 2010 to 2024 indicate that animal epidemics are highly spatially concentrated, primarily distributed in active foci involving species such as *Marmota himalayana, Marmota baibacina, Spermophilus undulatus, and Meriones unguiculatus*. The active periods of animal epidemics in these regions demonstrate spatiotemporal consistency with human case reports, suggesting that animal epidemics could serve as leading indicators of human plague risk. In terms of detection methods, the positivity rate of traditional pathogen isolation remains consistently low, while immunological assays yield persistently positive results, indicating widespread previous and latent infections among host animals. Given the large host population and complex host-vector network, once host density or climatic conditions exceed a certain threshold, dormant plague foci may be reactivated (*15*). Therefore, long-term systematic surveillance should be conducted in key areas (*16*–*17*). No single technology can simultaneously achieve speed, sensitivity, and confirmation. Therefore, a technological portfolio of "rapid PCR screening, immunological interpretation, and pathogen isolation confirmation" is recommended, supported by standardized quality control to ensure cross-regional comparability.

Risk Identification and Control Measures

The current situation for plague prevention and

control in China is complex, and the potential public health threats it poses cannot be ignored. To address these challenges, the prevention and control system can be optimized based on the following aspects:

Surveillance Precision: Identify core risk areas to enable precise resource allocation; Early Warning Technology: Build intelligent early warning models using multi-source data to enhance the capability for early epidemic identification; Prevention and Control Focus: Continuously strengthen basic prevention and control capacity in high-incidence areas, and emergency response capabilities for plague; Pre-positioned Resources: Proactively pre-position medical supplies and strengthen specialized personnel training before the epidemic season to reduce outbreak response start-up time; Localized Treatment: Establish a "treatment on-site" support system to reduce the risk of epidemic spread caused by cross-regional medical seeking.

Sample Size Limitations: Only 33 human cases were reported nationwide during the study period; Spatial Heterogeneity of the Surveillance Network: Surveillance site distribution may be uneven due to practical factors, including historical disease activity, geographical accessibility, and prevention and control resource allocation; Feasibility of Technology Promotion: High cost and stringent requirements of PCR and molecular techniques may limit their standardized application at the grassroots level.

International Comparison and Implications for Prevention and Control Strategies

From an international research perspective, plague has stable natural focus distributions in a few countries and regions across the Americas, Asia, and Africa (*1*,*18*). International studies have indicated that ecological factors, such as climate warming, changes in precipitation patterns, and vegetation "greening," can promote the activation and spatial expansion of these foci by influencing host and vector population dynamics (*16*,*19*). This study found that newly identified or marginally active foci in China were predominantly distributed along climate-ecological transition zones, a finding consistent with international research. In terms of surveillance and prevention, the integration of molecular detection techniques, such as PCR, with molecular typing and molecular tracing methods has become an international trend (*20*–*21*).

Surveillance data from 2010 to 2024 indicate that the plague situation in China remains complex, characterized by "sporadic human cases, active animal epidemics in some foci, and slow expansion of natural foci." To address these challenges, future prevention and control efforts should establish a tiered technological system of "PCR rapid screening, immunological trend analysis, and pathogen isolation for confirmation and traceability." This should be coupled with strengthened cross-provincial/regional collaboration and rapid data sharing, thereby systematically enhancing early warning and enabling precise responses, and continuously consolidating prevention and control achievements.

Conflicts of interest: No conflicts of interest.

doi: 10.46234/ccdcw2026.074

[#] Corresponding author: Kuidong Shao, shaokuidong001@163.com.

[1] Plague and Brucellosis Prevention and Control Base, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Baicheng City, Jilin Province, China.
[&] Joint first authors.

Submitted: November 27, 2025
Accepted: March 31, 2026
Issued: April 17, 2026

REFERENCES

1. World Health Organization. Plague—Fact sheet. 2022. https://www.who.int/news-room/fact-sheets/detail/plague. [2025-09-15].
2. Chinese Center for Disease Control and Prevention. Plague in China. Beijing: People's Medical Publishing House, 2019.
3. Barbieri R, Signoli M, Chevé D, Costedoat C, Tzortzis S, Aboudharam et al. Yersinia pestis: the natural history of plague. Clin Microbiol Rev 2020;34(1):e00044 – 19. https://doi.org/10.70675/360f36d1ze4f8z4fd5zb5cfz8994abe988d2.
4. Li H, Yao XH, Xu C, Cong XB, Ju C, Shao KD. Plague risk assessment — China, 2022. China CDC Wkly 2022;4(20):417 – 20. https://doi.org/10.46234/ccdcw2022.096.
5. National Administration of Disease Control and Prevention, National Health Commission of the People's Republic of China. National plague surveillance program (2024 revision). https://www.ndcpa.gov.cn. (In Chinese).
6. Sun Z, Xu L, Schmid BV, Dean KR, Zhang ZB, Xie Y, et al. Human plague system associated with rodent diversity and other environmental factors. R Soc Open Sci 2019;6(6):190216. https://doi.org/10.1098/rsos.190216.
7. Ministry of Health of the People's Republic of China. WS 279-2008 Diagnostic criteria for plague. Beijing: China Standards Press, 2008. (In Chinese).
8. Li XY, Liu SM, Shi XJ, Wang CY, Li L, Liu SY, et al. Impacts of climate change in China: northward migration of isohyets and reduction in cropland. Land 2025;14(7):1417. https://doi.org/10.3390/land14071417.
9. Xu L, Liu QY, Stige LC, Ben Ari T, Fang XY, Chan KS, et al. Nonlinear effect of climate on plague during the third pandemic in

China. Proc Natl Acad Sci USA 2011;108(25):10214 – 9. https://doi.org/10.1073/pnas.1019486108.
10. Xu L, Wang Q, Yang RF, Ganbold D, Tsogbadrakh N, Dong KX, et al. Climate-driven marmot-plague dynamics in Mongolia and China. Sci Rep 2023;13:11906. https://doi.org/10.1038/s41598-023-38966-1.
11. Ben-Ari T, Neerinckx S, Agier L, Cazelles B, Xu L, Zhang ZB, et al. Identification of Chinese plague foci from long-term epidemiological data. Proc Natl Acad Sci USA 2012;109(21):8196 – 201. https://doi.org/10.1073/pnas.1110585109.
12. Liu BX, Zhang DY, Chen YH, He ZK, Liu J, Lyu DY, et al. Epidemiological characteristics of plague in the *Meriones unguiculatus* plague focus — Inner Mongolia Autonomous Region, China, 1950-2019. China CDC Wkly 2020;2(49):935 – 45. https://doi.org/10.46234/ccdcw2020.256.
13. Moore SM, Monaghan A, Griffith KS, Apangu T, Mead PS, Eisen RJ. Improvement of disease prediction and modeling through the use of meteorological ensembles: human plague in Uganda. PLoS One 2012;7(9):e44431. https://doi.org/10.1371/journal.pone.0044431.
14. Li JY, Wang YM, Liu F, Shen XN, Wang YT, Fan MG, et al. Genetic source tracking of human plague cases in Inner Mongolia–Beijing, 2019. PLoS Negl Trop Dis 2021;15(8):e0009558. https://doi.org/10.1371/journal.pntd.0009558.
15. Qin JL, Wu YR, Shi LY, Zuo XJ, Zhang XLL, Qian XW, et al. Genomic diversity of *Yersinia pestis* from Yunnan Province, China, implies a potential common ancestor as the source of two plague epidemics. Commun Biol 2023;6(1):847. https://doi.org/10.1038/s42003-023-05186-2.
16. Ben-Ari T, Neerinckx S, Gage KL, Kreppel K, Laudisoit A, Leirs H, et al. Plague and climate: scales matter. PLoS Pathog 2011;7(9):e1002160. https://doi.org/10.1371/journal.ppat.1002160.
17. Li RY, Su C, Lou Z, Song ZZ, Pu EN, Li YQ, et al. Associations between ecological diversity and rodent plague circulation in Yunnan Province, China, 1983-2020: a data-informed modelling study. PLoS Negl Trop Dis 2023;17(6):e0011317. https://doi.org/10.1371/journal.pntd.0011317.
18. Randremanana R, Andrianaivoarimanana V, Nikolay B, Ramasindrazana B, Paireau J, Bosch QAT, et al. Epidemiological characteristics of an urban plague epidemic in Madagascar, August-November, 2017: an outbreak report. Lancet Infect Dis 2019;19(5):537 – 45. https://doi.org/10.1016/S1473-3099(18)30730-8.
19. Kausrud KL, Begon M, Ari TB, Viljugrein H, Esper J, Büntgen U, et al. Modeling the epidemiological history of plague in Central Asia: palaeoclimatic forcing on a disease system over the past millennium. BMC Biol 2010;8(1):112. https://doi.org/10.1186/1741-7007-8-112.
20. Loïez C, Herwegh S, Wallet F, Armand S, Guinet F, Courcol RJ. Detection of *Yersinia pestis* in sputum by real-time PCR. J Clin Microbiol 2003;41(10):4873 – 5. https://doi.org/10.1128/JCM.41.10.4873-4875.2003.
21. Vogler AJ, Keim P, Wagner DM. A review of methods for subtyping *Yersinia pestis*: from phenotypes to whole genome sequencing. Infect, Genet Evol 2016;37:21 – 36. https://doi.org/10.1016/j.meegid.2015.10.024.

Vital Surveillances

Genetic Analysis of Coxsackievirus A4 Among Healthy Children — Xizang Autonomous Region, China, 1996–2024

Kaitao Xiao[1,2,&]; Mei Hong[3,&]; Mengyi Xiao[1,2]; Fan Li[4]; Kexin Shao[1,2]; Chenglin Zhu[1,2]; Lan Yang[2]; Jie Lin[2]; Qin Wang[3]; Cidan Zhuoga[3]; Shuangli Zhu[2]; Dongmei Yan[2]; Yong Zhang[2]; Xiaomei Li[1,#]; Jinbo Xiao[2,#]

ABSTRACT

Introduction: Coxsackievirus A4 (CVA4) causes several diseases, including hand, foot, and mouth disease (HFMD) and herpangina. This study analyzed CVA4 isolates collected between 1996 and 2024 from the Xizang Autonomous Region to elucidate the phylogenetic characteristics and epidemiological patterns of this virus in high-altitude regions.

Methods: VP1 coding region sequences of CVA4 were obtained through virus isolation and Sanger sequencing. MEGA software was used to construct a maximum likelihood phylogenetic tree based on the VP1 region. The BEAST toolkit was used to generate a maximum clade credibility tree and perform phylogeographic analysis.

Results: From 2018 onwards, the prevalence of CVA4 among healthy carriers increased significantly, accounting for 62.12% of all detections. Genotyping revealed that most isolates belonged to genotype C, while the remainder were classified as D2. As the dominant genotype, genotype C has spread outward from Xigaze and Lhasa since 2011, leading to multiple asymptomatic infections in Shannan (2020), Xigaze and Lhasa (2023), and Ngari Prefecture (2024).

Conclusions: Transmission of CVA4 genotype C among healthy children in high-altitude areas suggests strong environmental adaptability, highlighting the need to strengthen enterovirus surveillance and control, including targeted monitoring of CVA4, in these regions.

Hand, foot, and mouth disease (HFMD) and herpangina (HA) are acute, highly contagious infectious diseases that predominantly affect children under five years of age. Coxsackievirus A4 (CVA4) is an important etiological agent of both conditions (*1*). As a member of the *Enterovirus* genus within the Picornaviridae family, CVA4 possesses a single-stranded positive-sense RNA genome of approximately 7.4 kb. The VP1 protein contains critical antigenic determinants involved in viral adsorption and host cell infection, and sequence variability in VP1 forms the basis for enterovirus serotype and genotype classification (*2*).

The prototype CVA4 strain (AY421762/High Point/1948) was first isolated in 1948 from urban sewage collected during a poliomyelitis outbreak in North Carolina, USA (*3*). Recent surveillance data indicate an increasing detection rate of CVA4 among patients with HFMD and HA, which correlates with its involvement in large-scale outbreaks across several Chinese provinces, including Jiangsu (*4*) and Beijing (*5*), as well as in other countries, such as Thailand (*6*) and Vietnam (*7*). CVA4 infection has been associated with severe neurological complications, suggesting that its pathogenic potential and epidemic risk may have been underestimated (*8*).

Previous studies have indicated that the prevalence of enteroviruses may be associated with climatic factors, such as temperature and humidity (*9*). The Xizang Autonomous Region has a unique climatic environment, and surveillance for CVA4 remains limited. Thus, the prevalence of CVA4 may have been underestimated. Clinical cases typically represent only the tip of the iceberg of viral transmission. Most transmission is driven by asymptomatic or paucisymptomatic infections among healthy children, particularly infants and young children, who serve as primary reservoirs and vectors for persistent viral circulation within communities (*10–11*). This study analyzed CVA4 isolates collected from healthy children in the Xizang Autonomous Region between 1996 and 2024 to elucidate the phylogenetic characteristics and epidemiological patterns of this virus in high-altitude regions.

METHODS

Sample Collection and Virus Identification

A total of 2,091 fecal specimens were collected from healthy children across seven prefecture-level divisions of the Xizang Autonomous Region: Ngari Prefecture, Qamdo, Lhasa, Nyingchi, Nagqu, Xigaze, and Shannan. All specimens were stored at 4 °C and transported to the National Laboratory for Poliomyelitis for further processing. Upon arrival, the samples were processed according to the Poliomyelitis Laboratory Manual (fourth edition), followed by viral isolation in human rhabdomyosarcoma and laryngeal epidermoid carcinoma cell lines.

Viral nucleic acids were extracted from the cell cultures using a Tianlong Ex-DNA/RNA Virus kit (CDC/T327, Tianlong, Xi'an) and the GeneRotex96 nucleic acid extraction instrument (Tianlong, Xi'an). The extracted nucleic acids were subjected to enterovirus-specific real-time reverse transcription polymerase chain reaction (real-time RT-PCR) using the One Step PrimeScript™ RT-PCR Kit (Perfect Real Time; TaKaRa, Dalian, China).

Molecular Typing of Enteroviruses and Sequencing of the CVA4 VP1 Coding Region

Real-time RT-PCR-positive samples were used to amplify the enterovirus VP1 coding region using the PrimeScript™ One Step RT-PCR Kit Ver.2 (TaKaRa, Dalian, China). The 25 μL reaction mixture consisted of 5 μL RNA template and 20 μL master mix, with primers E486/E488 (Enterovirus alpha-coxsackie universal primers), E490/E492 (Enterovirus beta-coxsackie universal primers), and E494/E496 (Enterovirus coxsackie-polio universal primers) (*12*). PCR products were sequenced by Sangon Biotech Co., Ltd. (Shanghai, China). The resulting AB1 files were assembled and edited using Sequencher software (version 5.4.5; GeneCode, Ann Arbor, MI, USA), and consensus sequences were subjected to a Basic Local Alignment Search Tool on the National Center for Biotechnology Information platform for enterovirus serotype determination.

Viral isolates confirmed as CVA4 were subjected to amplification of the full-length VP1 region (915 bp) using the primers forward CVA4-VP1-F: 5'-ACA CGCCGAACGAAGCTAAT-3' and reverse CVA4-VP1-R: 5'-TTATGTGTGGCTAGATGGCG-3'. Amplification was performed as previously described (*13*), and the full-length VP1 sequences of all CVA4 strains were assembled using Sequencher software.

Bioinformatic Analysis

The VP1 sequences obtained and 27 reference strains representing various CVA4 genotypes (*5*) were aligned using MEGA (version 12.1, freely available) (*14*). A maximum likelihood (ML) phylogenetic tree was constructed in MEGA using the optimal model identified by jModelTest (version 2.1.10, freely available), with branch support assessed via 1,000 bootstrap replicates. Genetic distances between sequences were calculated using MEGA and visualized as a heatmap using the corrplot package in R software (version 4.0.5; R Foundation for Statistical Computing, Vienna, Austria).

The phylogenetic tree was analyzed using TempEst software (version 1.5.3, freely available) via root-to-tip regression to assess temporal signals. Aligned sequences were imported into BEAUti, specifying a strict molecular clock and Bayesian SkyGrid coalescence model. The MCMC chain length was set to 1,000,000,000 generations, and a BEAST control file was generated. Phylodynamic inferences were performed using BEAST X software (version 10.5.0, freely available) (*15*), with convergence confirmed using Tracer software (v1.7.1; effective sample size >200). A maximum clade credibility (MCC) tree was generated using TreeAnnotator (10% burn-in) and visualized using FigTree. SpreaD3 (version 0.9.7, freely available) was used to calculate Bayes factors (BFs) (*16*), and CVA4 migration routes within the Xizang Autonomous Region were reconstructed based on BF values (BF≥3, posterior probability >0.5).

RESULTS

Demographic Characteristics of CVA4-positive Healthy Carriers

This study analyzed 66 asymptomatic CVA4 carriers identified through surveillance by the Xizang Autonomous Region Center for Disease Control and Prevention between 1996 and 2024 (Figure 1A). The number of carriers exceeded five in 6 years (2000, 2011, 2018, 2020, 2023, and 2024), with 41 cases (62.12%) reported since 2018. Among them, 25 were male and 38 were female; sex information was missing for three cases (Figure 1B). Most carriers were children

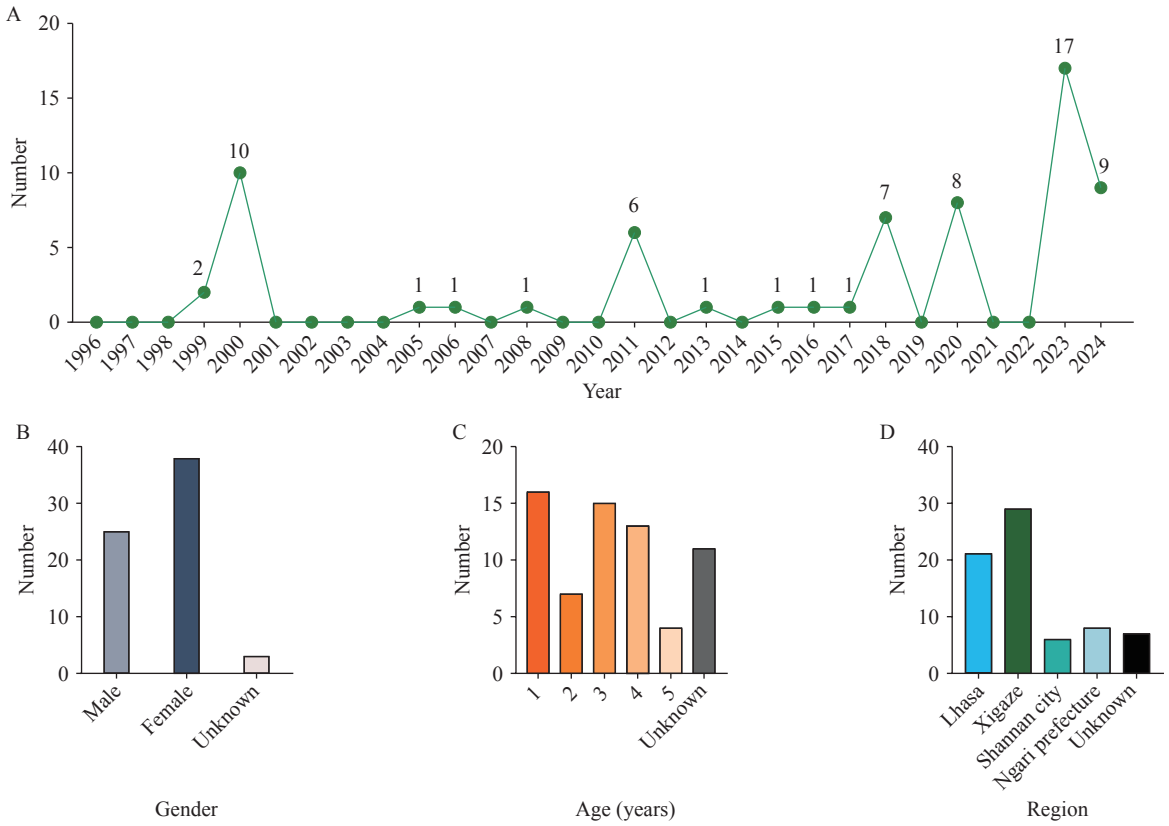


FIGURE 1. Demographic characteristics of CVA4-positive healthy carriers. (A) Annual detection count of CVA4 healthy carriers. (B) Sex distribution of carriers. (C) Age distribution of carriers. (D) Geographic distribution of carriers.
Abbreviation: CVA4=Coxsackievirus A4.

aged 0–5 years (Figure 1C), predominantly from Lhasa and Xigaze, with more than 20 cases identified in each region (Figure 1D).

Phylogenetic Analysis Based on CVA4 Full-length VP1 Sequences

A phylogenetic tree was constructed based on the full-length VP1 sequences from the 66 CVA4 strains obtained in this study and 27 reference strains (Figure 2A). CVA4 segregates into five genotypes (A–E), with genotype D split into sub-genotypes D1 and D2. Genotype A represents the prototype strain (AY421762/High Point/1948); B corresponds to the 1999 Kenyan strain; C consists entirely of Chinese isolates; D1 includes Chinese and Russian strains; D2 comprises Indian, U.S., and Chinese strains; and E contains Chinese and Russian strains. In the Xizang Autonomous Region, genotype C was predominant (56/66, 84.85%) followed by genotype D2 (10/66, 15.15%). Genotype C has been circulating continuously since it was first detected in 1999. The 56 genotype C strains formed two distinct phylogenetic branches: branch 1 contained mainly the 1999–2000 strains, whereas branch 2 consisted almost entirely of strains isolated after 2011. A heatmap of genetic distances confirmed a marked increase between the 2011–2024 sequences and earlier sequences (Figure 2B). Genotype D2 was first identified in Xizang in 2011, with sporadic cases reported in 2015 and 2017, culminating in a clustered outbreak in 2018.

Geographical Transmission Pattern of CVA4

Analysis of the MCC tree showed strong geographical clustering of CVA4 sequences. In the Xizang Autonomous Region, the virus diverged into genotypes C and D around 1977, consistent with the ML tree (Figure 3A). Genotype C splits into two branches. Branch 1 contained early strains that circulated mainly in Lhasa and Xigaze, with inferred transmission from Lhasa to Xigaze from 1997 to 1999. Branch 2 consisted of strains isolated after 2011, showing repeated cycles of transmission between Lhasa and Xigaze. Subsequent transmissions included Xigaze to Shannan (2018–2020), which led to a local cluster in 2020; Xigaze to Lhasa (2021–2023); and Lhasa to

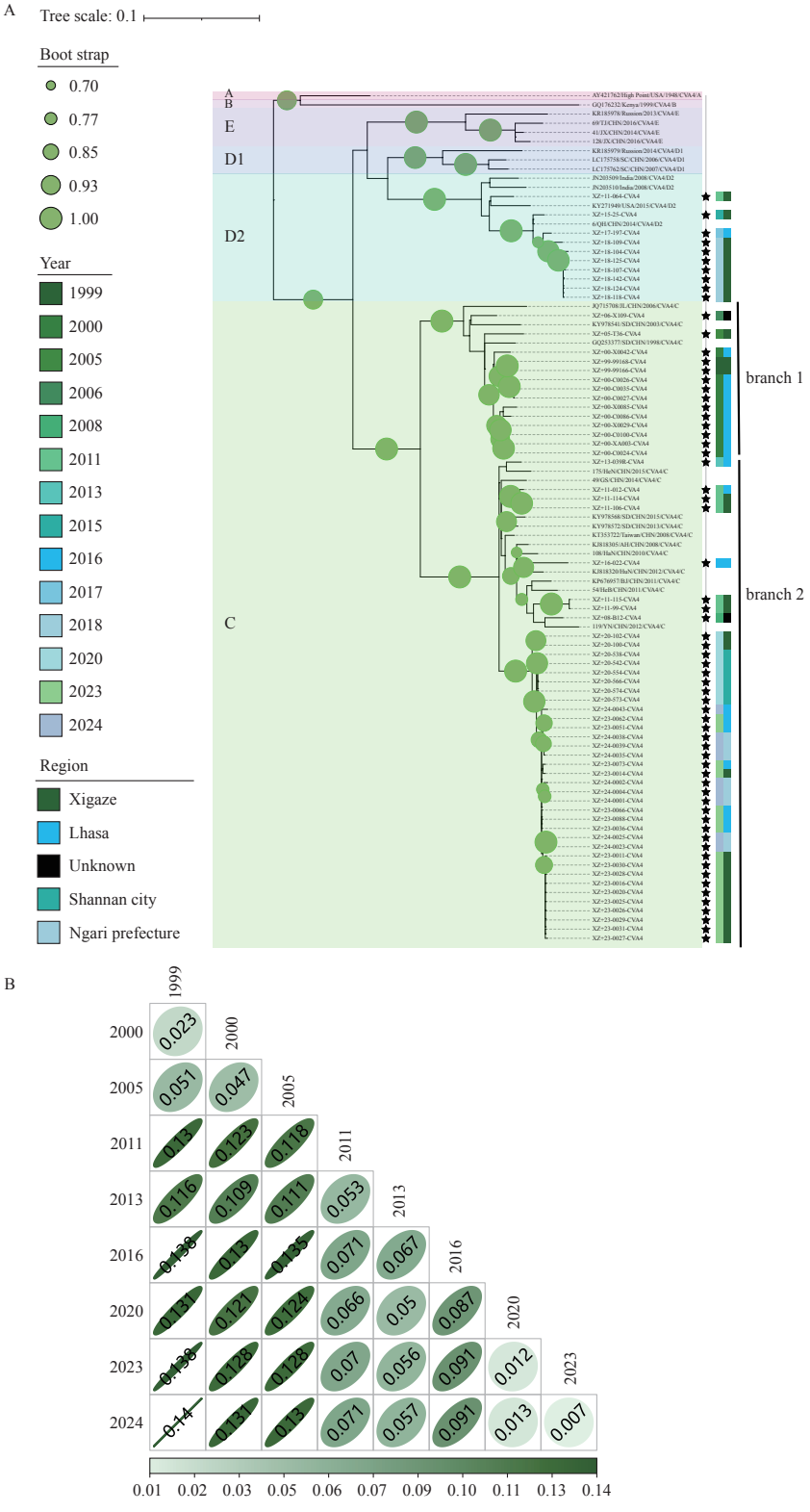


FIGURE 2. Phylogenetic analysis of Coxsackievirus A4. (A) Maximum likelihood phylogenetic tree of 66 Xizang CVA4 strains and 27 reference strains. (B) Heatmap of genetic distances among genotype C sequences grouped by collection year.

Note: For panel A, Genotypes A–E are indicated by colored backgrounds; stars denote Xizang strains. The first color-coded block shows collection year; the second block shows sampling region.

Abbreviation: CVA4=Coxsackievirus A4.

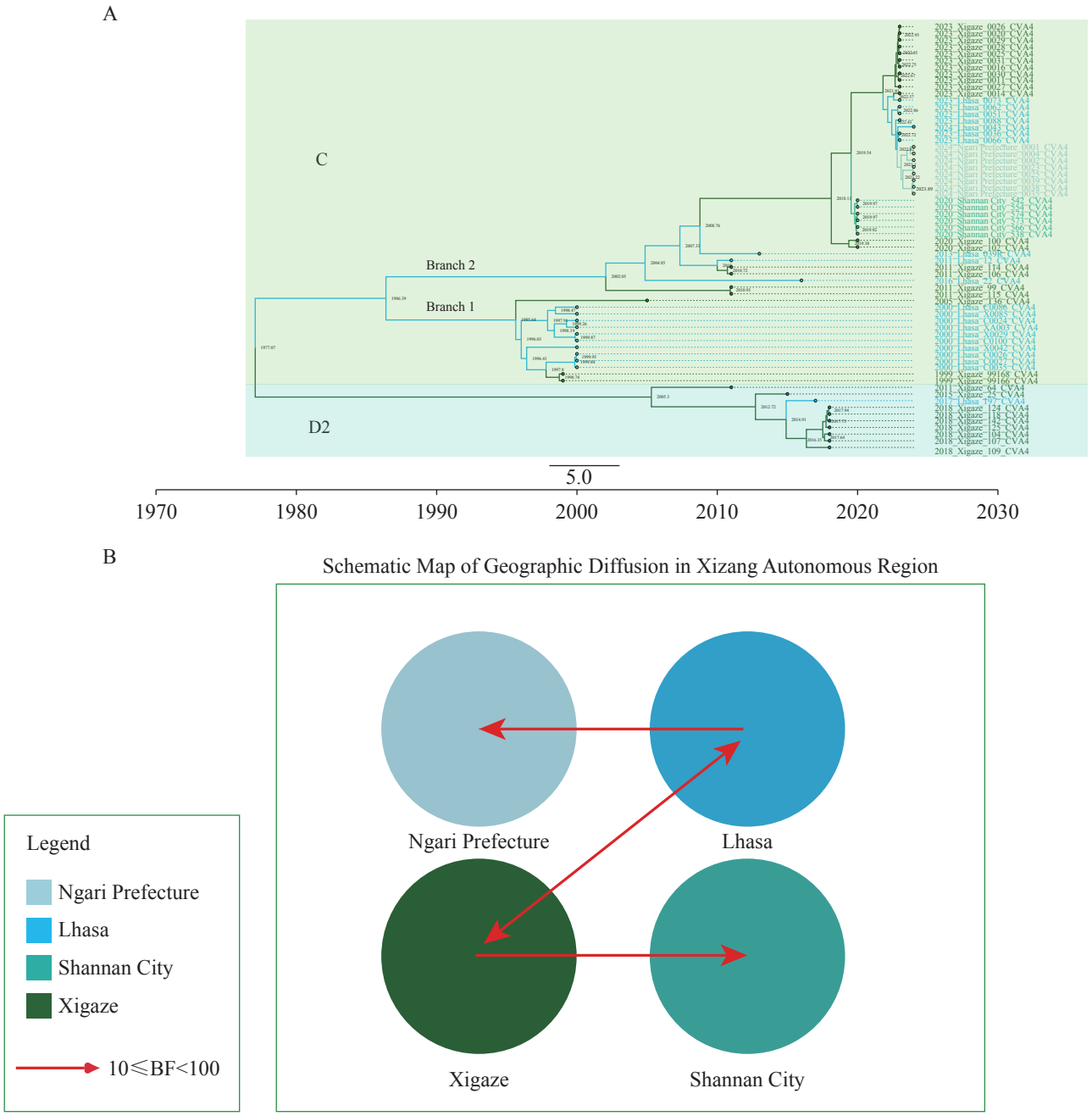


FIGURE 3. Geographical transmission pattern of CVA4. (A) Maximum clade credibility tree of 64 Xizang CVA4 sequences. (B) Transmission pathways inferred from SpreaD3 analysis; paths shown have pp>0.5 and BF≥3.
Abbreviation: CVA4=Coxsackievirus A4; pp=posterior probability; MCC=maximum clade credibility; BF=Bayes factor.

Ngari Prefecture (2022–2024). The D2 sub-genotype circulated only between Lhasa and Xigaze until its disappearance in 2018.

In the SpreaD3 analysis, transmission routes with a posterior probability (pp) >0.5 and a BF≥3 were defined as well-supported. Based on these criteria, four credible transmission pathways were identified (Figure 3B).

DISCUSSION

As an important member of the enterovirus alphacoxsackie family, CVA4 is one of the most frequently detected coxsackievirus types in non-polio enterovirus surveillance (*17*). The synthesis of long-term CVA4 surveillance data from the Xizang Autonomous Region is crucial for clarifying local

epidemiological patterns. Based on VP1 sequence analysis, this study confirmed that genotypes C and D were the dominant genotypes circulating in the Xizang Autonomous Region from 1996 to 2024, consistent with major CVA4 genotypes prevalent in other regions of China (*13*,*18*). The disappearance of the D2 sub-genotype during transmission suggests that genotype C may be advantageous in terms of viral fitness, transmissibility, or immune evasion. These advantages have enabled it to outcompete other genotypes and establish a stable lineage. Consequently, continuous monitoring is essential for detecting future shifts in dominant genotypes.

This analysis revealed a significant increase in the genetic distance between genotype C sequences before and after 2011. Following the 2000 outbreak, detection rates remained very low for approximately 10 years, indicating a possible surveillance gap. The 2011 epidemic strain may not be a direct descendant of earlier local strains but rather a newly introduced external variant, which could explain the observed phylogenetic divergence. Notably, the CVA4 C genotype circulating since 2011 has formed three distinct temporal clusters (2020, 2023, and 2024) on the MCC tree with short intra-cluster branches, indicating multiple localized outbreaks across the Xizang Autonomous Region. All inferred transmission routes are strongly supported by BF values (*19*).

As both a recipient of the virus from Xigaze and a source of its spread to Ngari, Lhasa has functioned as a key transmission hub, reflecting its central political and transportation roles. Repeated outbreaks underscore the considerable transmissibility of the post-2011 CVA4 C genotype among healthy children and highlight the need for continued surveillance. Despite the cold, high-altitude setting and sparse population, the CVA4 C genotype achieved sustained transmission, producing several discrete temporal clusters. This suggests that environmental factors alone do not halt viral spread, and that population mobility and child congregation settings are the principal drivers of outbreaks. Although this study did not analyze the full genomic characteristics of CVA4, it emphasizes the need to strengthen routine preventive measures and establish a molecular surveillance-based early warning system to reduce the risk of future large-scale CVA4 epidemics driven by asymptomatic transmission.

This study conducted surveillance among healthy children and found an increasing trend in CVA4 detection rates, indicating that the virus poses an epidemic risk in the Xizang Autonomous Region. Major transportation hubs are at particular risk of sustained transmission. These findings highlight the need to strengthen local surveillance systems and implement proactive intervention measures to prevent and control CVA4 spread. CVA4 surveillance should be incorporated into the national notifiable disease surveillance system for patients with HFMD to enable timely detection and response.

Conflicts of interest: No conflicts of interest.

Ethical statement: Approved by the Ethical Committee for Life Science and Medical Research Involving Human Subjects, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention (Approval No: IVDC2024-012).

Funding: Supported by the National Disease Prevention and Control Bureau Public Health Talent Training Support Project (Project Number: GJJKJ-2024-ZY); the National Key Laboratory for Infectious Disease Traceability, Early Warning, and Intelligent Decision-making (NITFID) funded project (Project Name: Research on Monitoring and Predictive Early Warning Technology for Unknown and Emerging Pathogens, Project Number: ZDGWNLJS25-36); and the National Key R&D Program Project (Project Name: Construction of Virus Surveillance Network Data Standards and Data Platform, Project Number: 2021YFC2302003).

doi: 10.46234/ccdcw2026.076

[#] Corresponding authors: Jinbo Xiao, xiaojb@ivdc.chinacdc.cn; Xiaomei Li, xmli@sdfmu.edu.cn.

[1] School of Public Health and Health Management, Shandong First Medical University (Shandong Academy of Medical Sciences), Jinan City, Shandong Province, China; [2] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases (NITFID), National Laboratory for Poliomyelitis, WHO WPRO Regional Polio Reference Laboratory, National Health Commission Key Laboratory of Microbial Genomics, National Health Commission Key Laboratory for Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China; [3] Department of Microbiological Testing, Tibet Autonomous Region Center for Disease Control and Prevention, Lhasa City, Xizang Autonomous Region, China; [4] National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China.
[&] Joint first authors.

Submitted: December 23, 2025
Accepted: March 08, 2026
Issued: April 17, 2026

REFERENCES

1. Wang DL, Yang QY, Zhu GY, Li ZN, Wu CY, Hu XJ, et al. Voxtalisib inhibits enterovirus 71 replication by downregulating host RAN and restoring IFN-STAT signaling. J Adv Res 2026;81:569 – 82. https://doi.org/10.1016/j.jare.2025.05.053.
2. Oberste MS, Maher K, Kilpatrick DR, Flemister MR, Brown BA, Pallansch MA. Typing of human enteroviruses by partial sequencing of VP1. J Clin Microbiol 1999;37(5):1288 – 93. https://doi.org/10.1128/JCM.37.5.1288-1293.1999.
3. Oberste MS, Peñaranda S, Maher K, Pallansch MA. Complete genome sequences of all members of the species *Human enterovirus A*. J Gen Virol 2004;85(Pt 6):1597-607. http://dx.doi.org/10.1099/vir.0.79789-0.
4. Zhang WT, Du YG, Gao X, Lin CQ, Hu R, Tong J. Pathogen spectra of hand, foot and mouth disease and herpangina in Xuzhou, Jiangsu, 2021. Dis Surveill 2024;39(7):863 – 8. https://doi.org/10.3784/jbjc.202307240375.
5. Wei XX, Xiao JB, Li M, Wang Y, Yang YN, Lv RR, et al. Genetic characteristics of coxsackievirus A4 associated with an epidemic of cluster fever. Chin J Virol 2025;41(5):1361 – 6. https://doi.org/10.13242/j.cnki.bingduxuebao.250117.
6. Puenpa J, Korkong S, Vichaiwattana P, Poovorawan Y. Genetic diversity and spread of recombinant coxsackievirus A4 in hand, foot, and mouth disease cases in Bangkok, Thailand: 2017-2023. Sci Rep 2024;14(1):26902. https://doi.org/10.1038/s41598-024-77832-6.
7. Hoa-Tran TN, Nguyen AT, Dao ATH, Kataoka C, Ta HTT, Nguyen HTV, et al. Genetic characterization of VP1 of coxsackieviruses A2, A4, and A10 associated with hand, foot, and mouth disease in Vietnam in 2012-2017: endemic circulation and emergence of new HFMD-causing lineages. Arch Virol 2020;165(4):823 – 34. https://doi.org/10.1007/s00705-020-04536-3.
8. Zhang ZJ, Zhang XC, Carr MJ, Zhou H, Li J, Liu SQ, et al. A neonatal murine model of coxsackievirus A4 infection for evaluation of vaccines and antiviral drugs. Emerg Microbes Infect 2019;8(1):1445 – 55. https://doi.org/10.1080/22221751.2019.1673135.
9. Pons-Salort M, Oberste MS, Pallansch MA, Abedi GR, Takahashi S, Grenfell BT, et al. The seasonality of nonpolio enteroviruses in the United States: patterns and drivers. Proc Natl Acad Sci USA 2018;115(12):3078 – 83. https://doi.org/10.1073/pnas.1721159115.
10. Li PY, Li T, Gu QY, Chen XM, Li JH, Chen XS, et al. Children's caregivers and public playgrounds: potential reservoirs of infection of hand-foot-and-mouth disease. Sci Rep 2016;6:36375. https://doi.org/10.1038/srep36375.
11. Jiao MMA, Apostol LN, de Quiroz-Castro M, Jee Y, Roque V Jr, Mapue II M, et al. Non-polio enteroviruses among healthy children in the Philippines. BMC Public Health 2020;20(1):167. https://doi.org/10.1186/s12889-020-8284-x.
12. Oberste MS, Maher K, Williams AJ, Dybdahl-Sissoko N, Brown BA, Gookin MS, et al. Species-specific RT-PCR amplification of human enteroviruses: a tool for rapid species identification of uncharacterized enteroviruses. J Gen Virol 2006;87(Pt 1):119-28. http://dx.doi.org/10.1099/vir.0.81179-0.
13. Ji TJ, Guo Y, Lv LK, Wang JX, Shi Y, Yu QL, et al. Emerging recombination of the C2 sub-genotype of HFMD-associated CV-A4 is persistently and extensively circulating in China. Sci Rep 2019;9(1):13668. https://doi.org/10.1038/s41598-019-49859-7.
14. Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K. MEGA12: molecular evolutionary genetic analysis version 12 for adaptive and green computing. Mol Biol Evol 2024;41(12):msae263. https://doi.org/10.1093/molbev/msae263.
15. Baele G, Ji X, Hassler GW, McCrone JT, Shao YC, Zhang ZY, et al. BEAST X for Bayesian phylogenetic, phylogeographic and phylodynamic inference. Nat Methods 2025;22(8):1653 – 6. https://doi.org/10.1038/s41592-025-02751-x.
16. Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. SpreaD3: interactive visualization of spatiotemporal history and trait evolutionary processes. Mol Biol Evol 2016;33(8):2167 – 9. https://doi.org/10.1093/molbev/msw082.
17. Fontana S, Buttinelli G, Fiore S, Amato C, Pataracchia M, Kota M, et al. Retrospective analysis of six years of acute flaccid paralysis surveillance and polio vaccine coverage reported by Italy, Serbia, Bosnia and Herzegovina, Montenegro, Bulgaria, Kosovo, Albania, North Macedonia, Malta, and Greece. Vaccines (Basel) 2021;10(1):44. https://doi.org/10.3390/vaccines10010044.
18. Wang M, Li J, Yao MX, Zhang YW, Hu T, Carr MJ, et al. Genome analysis of coxsackievirus A4 isolates from hand, foot, and mouth disease cases in Shandong, China. Front Microbiol 2019;10:1001. https://doi.org/10.3389/fmicb.2019.01001.
19. Kurdi B, Morris A, Cushman FA. The role of causal structure in implicit evaluation. Cognition 2022;225:105116. https://doi.org/10.1016/j.cognition.2022.105116.

Preplanned Studies

Epidemiological Characteristics and Transmissibility of Human Immunodeficiency Virus — Fujian Province, China, 1987–2024

Mingya Zhang[1,&]; Yanshu Ke[2,&]; Yue He[3,&]; Qiaoling Lian[1]; Chunyang Zhang[1]; Shaobin Wu[1]; Zhenghua Wang[1]; Xiaoli Lu[1]; Meirong Xie[1]; Shouli Wu[1]; Jianfeng Xie[1]; Zeyu Zhao[2]; Jia Rui[4]; Tianmu Chen[2]; Yuefeng Qiu[1,#]

Summary

What is already known about this topic?

Given the proportion of undiagnosed human immunodeficiency virus (HIV) cases and regional disparities in Fujian Province, achieving the 2030 targets to end acquired immunodeficiency syndrome (AIDS) remains uncertain.

What is added by this report?

HIV transmission in Fujian has stabilized through treatment and viral suppression, achieving the "95-95-95" targets; however, a diagnostic gap persists (86.30%). Transmissibility recently increased (R_{eff}=1.03), with males and adults aged 20–70 years as the primary high-risk populations, threatening elimination goals.

What are the implications for public health practice?

Improving diagnostic coverage is essential for epidemic elimination. Targeted screening and self-testing must be intensified among key populations, particularly middle-aged and older adults, while maintaining treatment achievements to prevent resurgence.

ABSTRACT

Introduction: Human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) remains a global public health threat, with distinct regional patterns across China. In this study, we examined the unique epidemic characteristics of Fujian Province to assess HIV transmission dynamics, evaluate progress toward the "95-95-95" targets, and optimize local prevention strategies.

Methods: HIV/AIDS surveillance data from Fujian Province spanning 1987–2024 were collected to analyze epidemiological characteristics and develop a comprehensive transmission dynamics model. Using the next-generation matrix method, we estimated the effective reproduction number (R_{eff}) to assess transmissibility. By integrating spatiotemporal distribution analysis with evaluation of progress toward the 95-95-95 targets, we assessed epidemic trends and intervention effectiveness.

Results: The HIV/AIDS epidemic in Fujian Province showed an initial increase followed by gradual stabilization, with notable spatiotemporal heterogeneity. High-risk populations included males (81.86%) and individuals aged 20–70 years (89.21%). Although treatment coverage and viral suppression achieved the 95% targets, diagnostic coverage (86.30%) remained suboptimal. Model estimates revealed that R_{eff} declined from 2–5 during 2009–2019 to below 1 after 2020, but rebounded to 1.03 by 2024.

Conclusion: Although HIV transmission in Fujian Province has stabilized, the persistent diagnostic gap and rebound in transmissibility in 2024 indicate potential resurgence risks. Future interventions should prioritize middle-aged and older adults, and enhanced case identification strategies are essential to support achievement of the 2030 HIV elimination targets.

Acquired immunodeficiency syndrome (AIDS), caused by human immunodeficiency virus (HIV), remains a major global public health challenge. By 2023, approximately 39.9 million people worldwide were living with HIV, with 1.3 million new infections and 630,000 AIDS-related deaths (*1*). China has a relatively low HIV/AIDS prevalence; however, significant regional disparities persist, with heterosexual transmission being the predominant route of infection. As of September 2018, China reported 850,000 people living with HIV and 262,000 AIDS-related deaths (*2*). Given the absence of a definitive cure for HIV infection, prevention strategies remain fundamental for epidemic control. In 2021, the United Nations adopted the Political Declaration on Ending AIDS 2030, and the World Health Organization released the Global Health Sector Strategy 2022–2030, establishing

the "95-95-95" targets: 95% of infected individuals knowing their status, 95% receiving antiretroviral therapy, and 95% achieving viral suppression by 2030 (*3*). By 2020, China had achieved 75%, 92%, and 96%, respectively, with undiagnosed cases representing the primary obstacle (*4*). Fujian Province has a relatively high prevalence of sexually transmitted infections (*5*), and increasing HIV incidence poses substantial challenges to meeting the 2030 elimination targets. Mathematical modeling provides valuable tools for assessing HIV transmission dynamics and predicting epidemic trends.

In this study, we constructed a comprehensive transmission dynamics model based on HIV disease progression, calculated the effective reproduction number (R_{eff}) using the next-generation matrix method, and evaluated the feasibility of eliminating AIDS in Fujian Province by 2030. Our findings provide an evidence-based foundation for optimizing regional HIV/AIDS prevention strategies.

We analyzed data from the HIV/AIDS reporting system of Fujian Province to examine epidemiologic characteristics from 1987 to 2024, focusing on spatiotemporal distribution and demographic trends.

We developed a transmission dynamics model based on HIV epidemiologic characteristics, following previous studies (*6*–*7*) (Supplementary Figure S1, available at https://weekly.chinacdc.cn/). The model included the following compartments: S (susceptible), I (undiagnosed HIV-infected), T_1 (diagnosed HIV-infected), T_2 (receiving antiretroviral therapy), T_3 (suppressed viral load), and A (people living with AIDS). Model assumptions were as follows: 1) Newborns are susceptible (S), with birth (b_r) and mortality rates (d_r). 2) Susceptible individuals (S) may become infected through contact with individuals in compartments I, A, T_1, or T_2. The baseline transmission rate for I is β, and the relative transmissibility of A, T_1, and T_2 is κ, κ_1, and κ_2 times that of I, respectively. 3) Parameter p_1 represents diagnostic coverage. The monthly transition rate from I to T_1 is $p_1 I$, whereas individuals in I progress to AIDS at a rate of $(1-p_1)\omega_1 I$. 4) Parameter p_2 represents treatment coverage. Individuals in T_1 compartment transition to T_2 at a monthly rate of $p_2 \delta T_1$, while progression to AIDS occurs at a rate of $(1-p_2)\omega_2 T_1$. 5) Parameter p_3 denotes the viral suppression rate. Individuals in T_2 progress to T_3 at a monthly rate of $p_3 \gamma T_2$, whereas progression to AIDS occurs at a rate of $(1-p_3)\omega_3 T_2$. 6) Parameter q represents the proportion of patients progressing to AIDS despite viral suppression. Individuals in T_3 transition to AIDS at a monthly rate of $q\omega_4 T_3$. 7) Disease-induced mortality rates are denoted by f_1 for HIV-infected individuals and f_2 for individuals with AIDS. Parameter details are provided in Table 1. The differential equation system is presented in Supplementary Figure S1B.

The basic reproduction number (R_0) represents the expected number of secondary infections generated by one infected individual in a completely susceptible population without interventions. The R_{eff} measures the transmission potential under ongoing intervention and was calculated using the next-generation matrix method (*8*). The detailed formula is provided in the Supplementary Materials.

Between 1987 and 2024, Fujian Province reported 30,602 HIV/AIDS cases and 6,032 deaths. The epidemic showed distinct phases: annual cases remained below 10 during 1987–1996, increased gradually during 1997–2003, surpassed 100 after 2004, exceeded 1,000 beginning in 2013, and peaked in 2019 before stabilizing (Figure 1A). Since 2010, incidence patterns have shown consistent troughs in February, with mortality mirroring these seasonal variations. Annual deaths exceeded 100 in 2009 and 500 in 2018. After 2008, the annual mortality rate consistently remained above 1 per 100,000 population (Figure 1B). The spatial distribution of HIV/AIDS exhibited pronounced regional disparities, with higher prevalence in coastal areas and lower prevalence inland. Since 2004, prevalence has shifted from localized clustering to province-wide distribution.

Among the 30,602 patients with HIV/AIDS, 25,050 (81.86%) were male and 5,552 (18.14%) were female, yielding a male-to-female ratio of 4.51∶1 (Figure 2A). Males consistently outnumbered females across all periods, comprising approximately 80% of cases since 2012. The predominant age range was 20–70 years, accounting for 27,300 patients (89.21%) (Figure 2B). In 1991, most cases occurred among individuals aged 20–40 years. By 2020, a bimodal distribution emerged, with peaks among young adults (20–40 years) and middle-aged to older adults (50–70 years). By 2024, the 50–70-year age group exhibited a higher incidence peak than the 20–40-year cohort, suggesting a shift in disease burden toward older populations.

Fujian Province showed substantial progress toward the "95-95-95" targets during 2009–2024 (Figure 3). Diagnostic coverage increased from 26.90% (2009) to 86.30% (2024), remaining below the 95% target

TABLE 1. Definitions and values of model parameters.

Parameters	Definition	Unit	Value	Range	Methods
β	Transmission rate coefficient of HIV-infected individuals, the effective contact rate per month	$Month^{-1}$	–	0–1	Model fitting
κ	Transmission rate coefficient of patients with AIDS	1	0.2692	0–1	Reference
κ_1	Transmission rate coefficient of individuals diagnosed with HIV	1	–	0–1	Model fitting
κ_2	Transmission rate coefficient of ART-treated individuals	1	0.1111	0–1	Reference
p_1	Diagnosis coverage among individuals infected with HIV	1	–	0.2690–0.8620	Data collection
p_2	Treatment coverage among individuals diagnosed with HIV	1	–	0.2452–0.9705	Data collection
p_3	Viral suppression rate among ART-treated individuals	1	–	0.8482–0.9656	Data collection
q	Proportion of virally suppressed individuals progressing to AIDS	1	0.0500	0–1	Hypothesis
ω_1	Progression rate from HIV infection to AIDS	$Month^{-1}$	0.0093	0.0064–0.0189	References
ω_2	Progression rate from HIV diagnosis to AIDS	$Month^{-1}$	0.0083	–	Hypothesis
ω_3	Progression rate from ART-treated individuals to AIDS	$Month^{-1}$	0.0076	–	Hypothesis
ω_4	Progression rate from viral suppression individuals to AIDS	$Month^{-1}$	0.0069	–	Hypothesis
δ	Progression rate from diagnosis to treatment	$Month^{-1}$	0.9709	0.0592– 3.3333	Reference
γ	Progression rate from treatment to viral suppression	$Month^{-1}$	0.1111	0.0833–0.1429	References
f_1	Fatality rate of individuals infected with HIV	$Month^{-1}$	–	0–0.1111	Data collection
f_2	Fatality rate of patients with AIDS	$Month^{-1}$	–	0–0.3333	Data collection
b_r	Birth rate	$Month^{-1}$	–	0.0006–0.0021	Statistical yearbook
d_r	Death rate	$Month^{-1}$	–	0.0004–0.0006	Statistical yearbook

Note: "–" means non-unique values, encompassing both year-to-year variations and changes across different time periods.
Abbreviation: HIV=human immunodeficiency virus; ART=antiretroviral therapy; AIDS=acquired immunodeficiency syndrome.

despite a temporary decline in 2018. Treatment coverage increased steadily, reaching 97.05% by 2024 and surpassing the 95% goal. Viral suppression rates exceeded 90% by 2011 and remained above 95% after 2019, achieving the target before 2030.

We fitted the HIV transmission dynamics model to reported case numbers from 2009 to 2024. Model fitting was conducted across three periods: the slow-growth phase (2009–2012), rapid-growth phase (2013–2019), and stabilization phase (2020–2024) (Supplementary Figure S2A, available at https://weekly.chinacdc.cn/). Goodness-of-fit analysis showed statistically significant results (R^2=0.864, P<0.001), indicating robust model performance (Supplementary Figure S2B).

The monthly R_{eff} of HIV in Fujian Province from 2009 to 2024 (Supplementary Figure S3A, available at https://weekly.chinacdc.cn/) revealed three phases: 1) a slow-growth period (2009–2012), with R_{eff} fluctuating consistently above the epidemic threshold; 2) a rapid-expansion phase (2013–2019), with R_{eff} values ranging from 2.0 to 4.0 above the threshold; and 3) a stabilization phase (2020–2024), during which R_{eff} declined toward 1.0, fell below the threshold after late 2022, and resurged to 1.03 by December 2024. Box plot analysis (Supplementary Figure S3B) confirmed these trends. From 2009 to 2019, R_{eff} exceeded 1.0 with considerable monthly variability, whereas by 2023 it fell below the epidemic threshold and became more tightly clustered, indicating improved epidemic control.

DISCUSSION

In this study, we analyzed HIV/AIDS epidemiological patterns in Fujian Province and developed a transmission dynamics model. Through model fitting, we calculated transmissibility parameters, providing a foundation for understanding epidemic dynamics and informing optimized control strategies.

The spatiotemporal evolution of HIV/AIDS in Fujian Province exhibited phase progression characterized by regional heterogeneity. From 2020 to 2024, the epidemic exhibited a gradual upward trend, with a markedly slower growth rate than that observed before 2019. Annual reporting troughs each February since 2010 coincided with Spring Festival migration, which may reflect the "Spring Festival effect" reported for certain sexually transmitted infections in previous studies (*9*). No clear seasonal variation was observed in

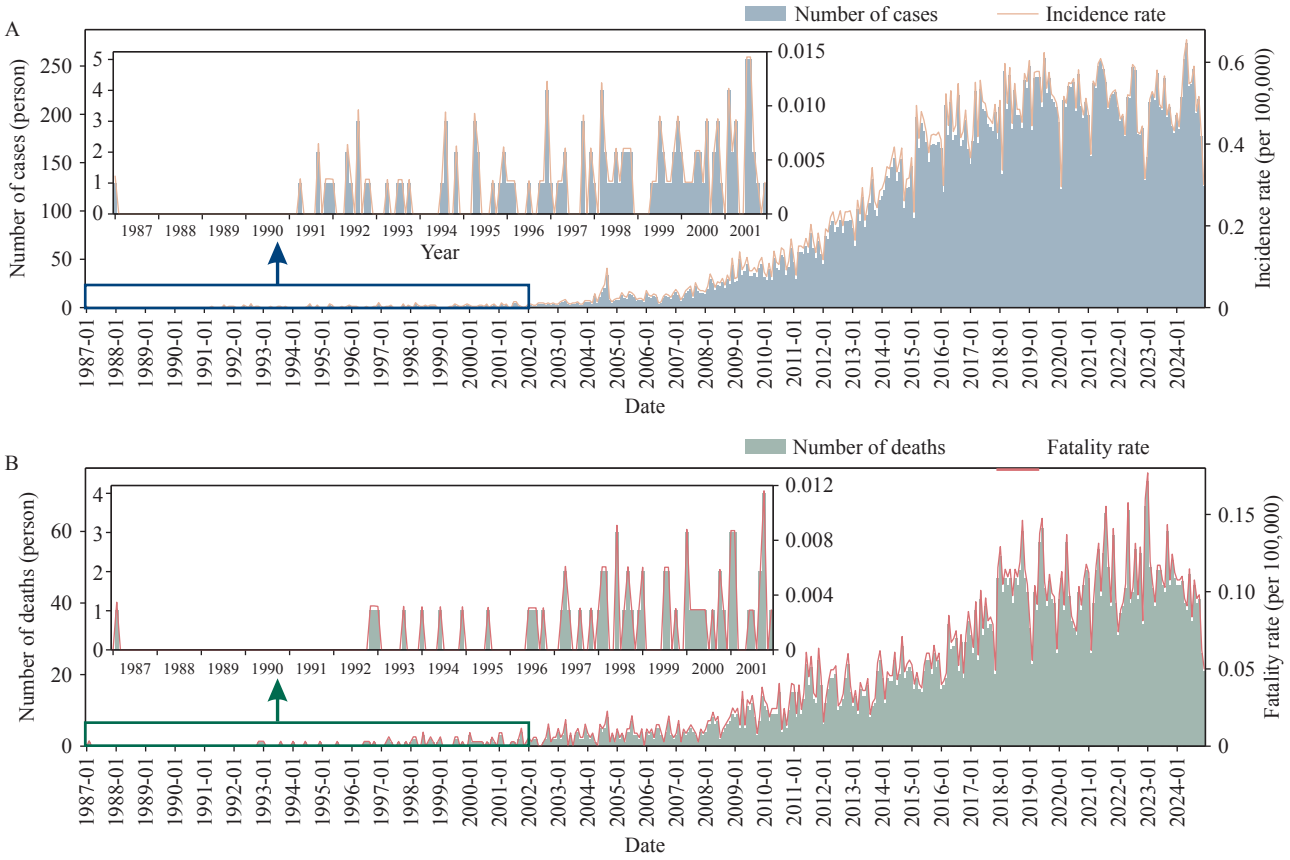


FIGURE 1. Temporal distribution of HIV/AIDS in Fujian Province, 1987–2024. (A) Reported HIV/AIDS cases and incidence rate (per 100,000 population),1987-2024; (B) AIDS-related deaths and case fatality rate (per 100,000 population), 1987-2024.
Abbreviation: HIV=human immunodeficiency virus; AIDS=acquired immunodeficiency syndrome.

HIV incidence. Spatial distribution correlates with economic development and population mobility. Males consistently accounted for 80% of cases after 2012, yielding a 4.51∶1 male-to-female ratio exceeding the national average and likely reflecting coastal migration patterns (*10*). Age distribution revealed a bimodal pattern, with middle-aged and older adults surpassing young adults, consistent with global demographic transitions associated with antiretroviral therapy benefits (*11*). Prevention strategies should emphasize pre-exposure prophylaxis (PrEP) implementation among younger populations while strengthening screening among older adults (*12*).

Fujian Province has made substantial progress toward the "95-95-95" targets, although indicators developed unevenly. Although diagnostic coverage exceeded national levels, it remained below the 2030 AIDS elimination target. Treatment coverage and viral suppression rates surpassed their targets. This disparity highlights the need to strengthen proactive screening among high-risk populations and enhance treatment adherence programs.

The epidemic exhibited distinct phases. During 2009–2012, R_{eff} fluctuated between 2.5–5, reflecting limited coverage of early prevention measures (*13–14*). We developed a comprehensive HIV transmission dynamics model for Fujian Province spanning 2009–2024, calibrated using a three-phase structure aligned with major shifts in provincial intervention strategies. The model demonstrated robust performance with strong statistical fit. The expansion of testing and treatment policies during 2013–2019 contributed to a decline in R_{eff}, although it remained above the epidemic threshold. A marked decline in R_{eff} was observed in early 2020, likely associated with public health and social measures implemented during the COVID-19 pandemic (*15*). From 2020 to 2024, R_{eff} remained lower than in the pre-2019 period, while case numbers increased at a substantially slower rate. The slight rebound in R_{eff} observed at the end of 2024 may reflect persistent undiagnosed infections. Strengthening case detection and sustaining high

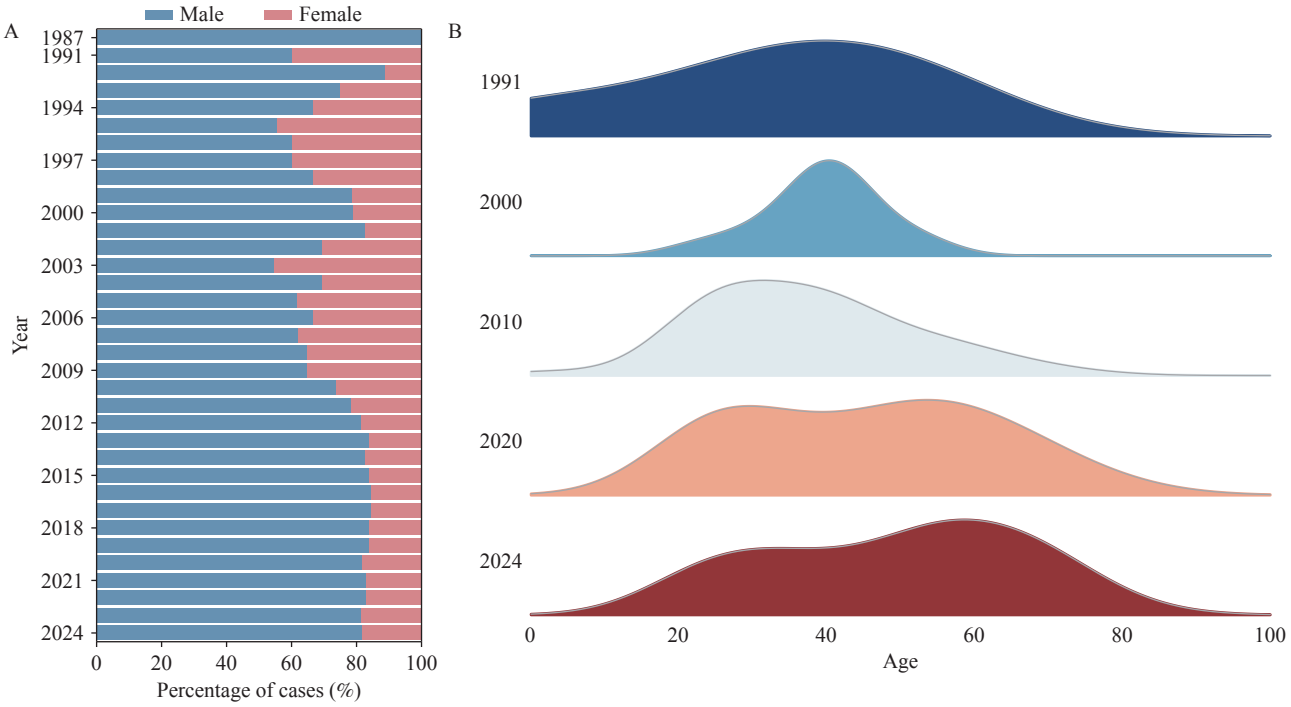


FIGURE 2. Demographic distribution of HIV/AIDS cases in Fujian Province, 1987–2024. (A) Sex ratio of reported HIV/AIDS cases from 1987 to 2024 (data unavailable during 1988–1990). (B) Age distribution of reported HIV/AIDS cases in 1991, 2000, 2010, 2020, and 2024.
Abbreviation: HIV=human immunodeficiency virus; AIDS=acquired immunodeficiency syndrome.

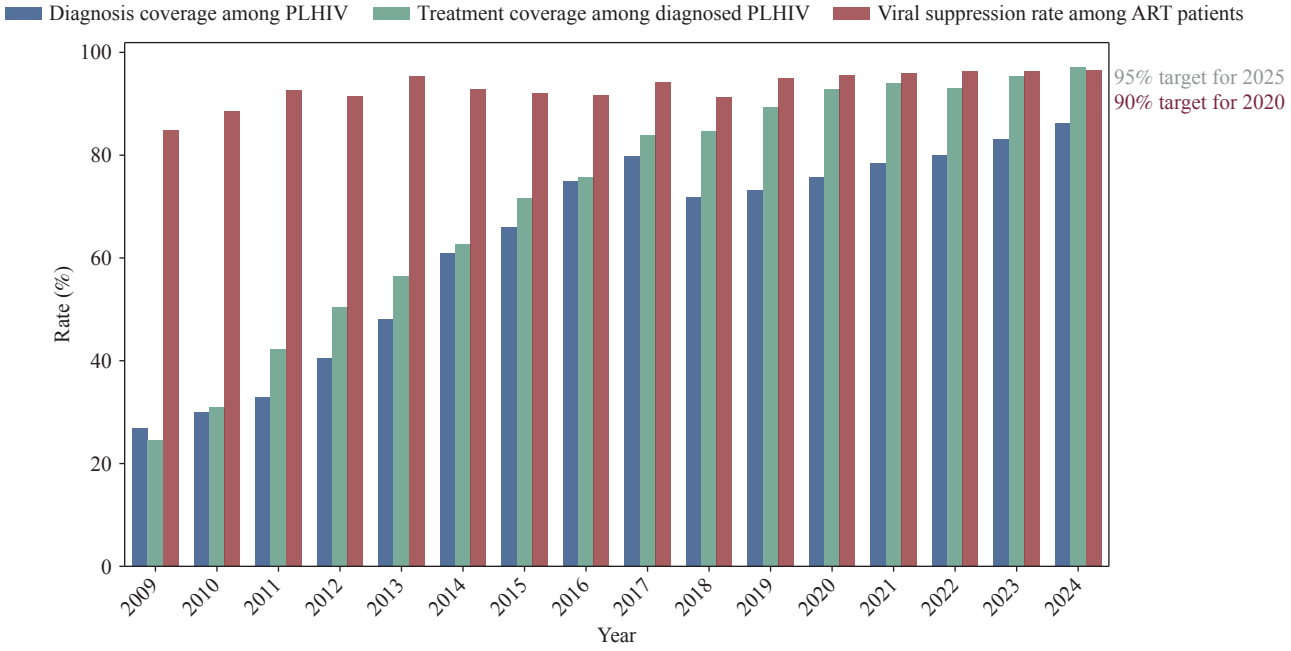


FIGURE 3. Progress toward the "95-95-95" targets in Fujian Province, 2009–2024.
Abbreviation: PLHIV=people living with HIV; ART=antiretroviral therapy.

treatment coverage remain essential to achieving the "95-95-95" targets by 2030.

This study has several limitations that require consideration. First, the transmission dynamics model treated the population as homogeneous and did not account for individual behavioral patterns or social network structures influencing HIV transmission. Second, the model did not simulate specific

intervention measures because current prevention strategies target high-risk populations through approaches that cannot be readily quantified using mathematical parameters. Future research should incorporate these behavioral and network factors to provide more comprehensive evaluations of progress toward ending the HIV epidemic by 2030.

In conclusion, the HIV/AIDS epidemic in Fujian Province demonstrated significant spatiotemporal heterogeneity, with coastal regions representing high-incidence areas. Males and individuals aged 20–70 years constituted the primary affected populations. The transmission dynamics model showed that R_{eff} consistently exceeded the epidemic threshold throughout 2009–2019. After 2020, transmissibility declined below unity, although an upward trajectory emerged in 2024, signaling potential resurgence risk. While treatment coverage and viral suppression achieved the 95% benchmark, diagnostic coverage remained suboptimal. Future prevention efforts must prioritize targeted interventions for middle-aged and older adults, expand self-testing accessibility, and address diagnostic gaps to realize the 2030 objective of ending the HIV epidemic.

doi: 10.46234/ccdcw2026.077

[#] Corresponding author: Yuefeng Qiu, qiuyuefeng@fjcdc.com.cn.

[1] Fujian Provincial Key Laboratory of Zoonosis Research, Fujian Provincial Center for Disease Control and Prevention, Fujian Provincial Academy of Preventive Medicine Fuzhou City, Fujian Province, China; [2] State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, National Innovation Platform for Industry-Education Integration in Vaccine Research, Xiamen University, Xiamen City, Fujian Province, China; [3] State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, Xiamen City, Fujian Province, China; [4] Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha City, Hunan Province, China.

[&] Joint first authors.

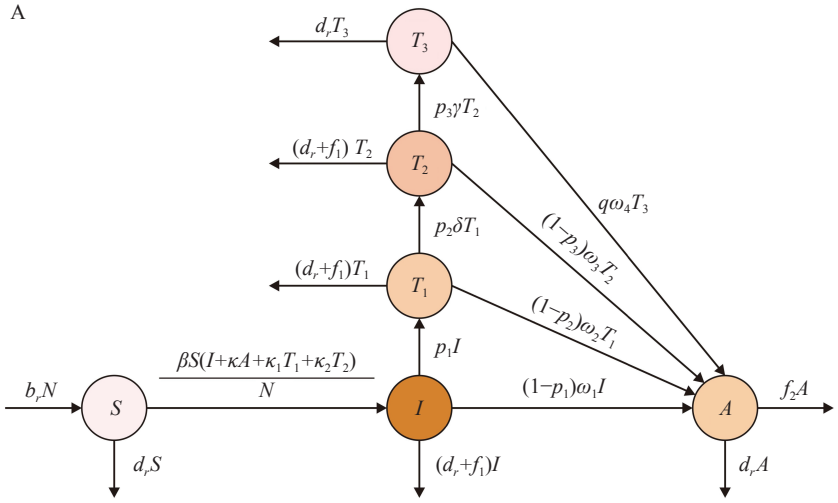
Submitted: October 23, 2025
Accepted: February 24, 2026
Issued: April 17, 2026

REFERENCES

1. UNAIDS. Global HIV & AIDS statistics — Fact sheet. 2024. https://www.unaids.org/en/resources/fact-sheet. [2026-02-16].
2. National Center for AIDS/STD Control and Preventive, China CDC. Progress in AIDS prevention and control work in our country 2018. 2018. https://ncaids.chinacdc.cn/xxgx/yqxx/201811/t20181123_197488.htm. (In Chinese). [2026-02-16].
3. UNAIDS. Political declaration on HIV and AIDS: ending inequalities and getting on track to end AIDS by 2030. 2021. https://www.unaids.org/sites/default/files/media/documents/2021_political-declaration-on-hiv-and-aids_zh.pdf. (In Chinese). [2026-02-16].
4. UNAIDS. 2020-2021 biennial report of the China-UN joint AIDS project China. 2021. http://hkb980dd.pic44.websiteonline.cn/upload/2020-2021BiennialReport-CN.pdf. (In Chinese). [2026-02-16].
5. Zheng JY, Zhang N, Shen GQ, Liang FC, Zhao Y, He XC, et al. Spatiotemporal and seasonal trends of class A and B notifiable infectious diseases in China: retrospective analysis. JMIR Public Health Surveill 2023;9:e42820. https://doi.org/10.2196/42820.
6. Guo ZY, Xiao D, Xu S, He K. Analysis and forecast of the HIV/AIDS epidemic in Mainland China, 1985-2016. J Public Health 2020;42(4):E458 – 67. https://doi.org/10.1093/pubmed/fdz116.
7. Li ZM, Teng ZD, Miao H. Modeling and control for HIV/AIDS transmission in China based on data from 2004 to 2016. Comput Math Methods Med 2017;2017:8935314. https://doi.org/10.1155/2017/8935314.
8. van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 2002;180(1-2):29 – 48. https://doi.org/10.1016/S0025-5564(02)00108-6.
9. Wei S, Lu YH, Gao MY, Wei GR, Jiang QW, Zhao NQ, et al. "Spring Festival Effects" on the main notifiable communicable diseases in China. Fudan Univ J Med Sci 2013;40(2):153 – 8. https://doi.org/10.3969/j.issn.1672-8467.2013.02.005.
10. Zha SL. Analysis of current situation of HIV epidemic in China in the Year 2012—2015. J Weinan Normal Univ 2016;31(8):25 – 30. https://doi.org/10.3969/j.issn.1009-5128.2016.08.005.
11. Kohli R, Klein RS, Schoenbaum EE, Anastos K, Minkoff H, Sacks HS. Aging and HIV infection. J Urban Health 2006;83(1):31 – 42. https://doi.org/10.1007/s11524-005-9005-6.
12. Torres TS, Luz PM, De Boni RB, de Vasconcellos MTL, Hoagland B, Garner A, et al. Factors associated with PrEP awareness according to age and willingness to use HIV prevention technologies: the 2017 online survey among MSM in Brazil. AIDS Care 2019;31(10):1193 – 202. https://doi.org/10.1080/09540121.2019.1619665.
13. Sun XH, Lu F, Wu ZY, Poundstone K, Zeng G, Xu P, et al. Evolution of information-driven HIV/AIDS policies in China. Int J Epidemiol 2010;39 Suppl 2(Suppl 2):ii4-13. http://dx.doi.org/10.1093/ije/dyq217.
14. Xie ZY, Duan ZZ. Evolving challenges in the implementation of China's "Four Frees and One Care" policy for HIV/AIDS. Risk Manag Healthc Policy 2024;17:1315 – 21. https://doi.org/10.2147/RMHP.S459173.
15. Booton RD, Fu GF, MacGregor L, Li JJ, Ong JJ, Tucker JD, et al. The impact of disruptions due to COVID-19 on HIV transmission and control among men who have sex with men in China. J Int AIDS Soc 2021;24(4):e25697. https://doi.org/10.1002/jia2.25697.

SUPPLEMENTARY MATERIAL

A



B

$$\frac{dS}{dt}=b_rN-\frac{\beta S(I+\kappa A+\kappa_1T_1+\kappa_2T_2)}{N}-d_rS$$

$$\frac{dI}{dt}=\frac{\beta S(I+\kappa A+\kappa_1T_1+\kappa_2T_2)}{N}-p_1I-(1-p_1)\omega_1I-(d_r+f_1)I$$

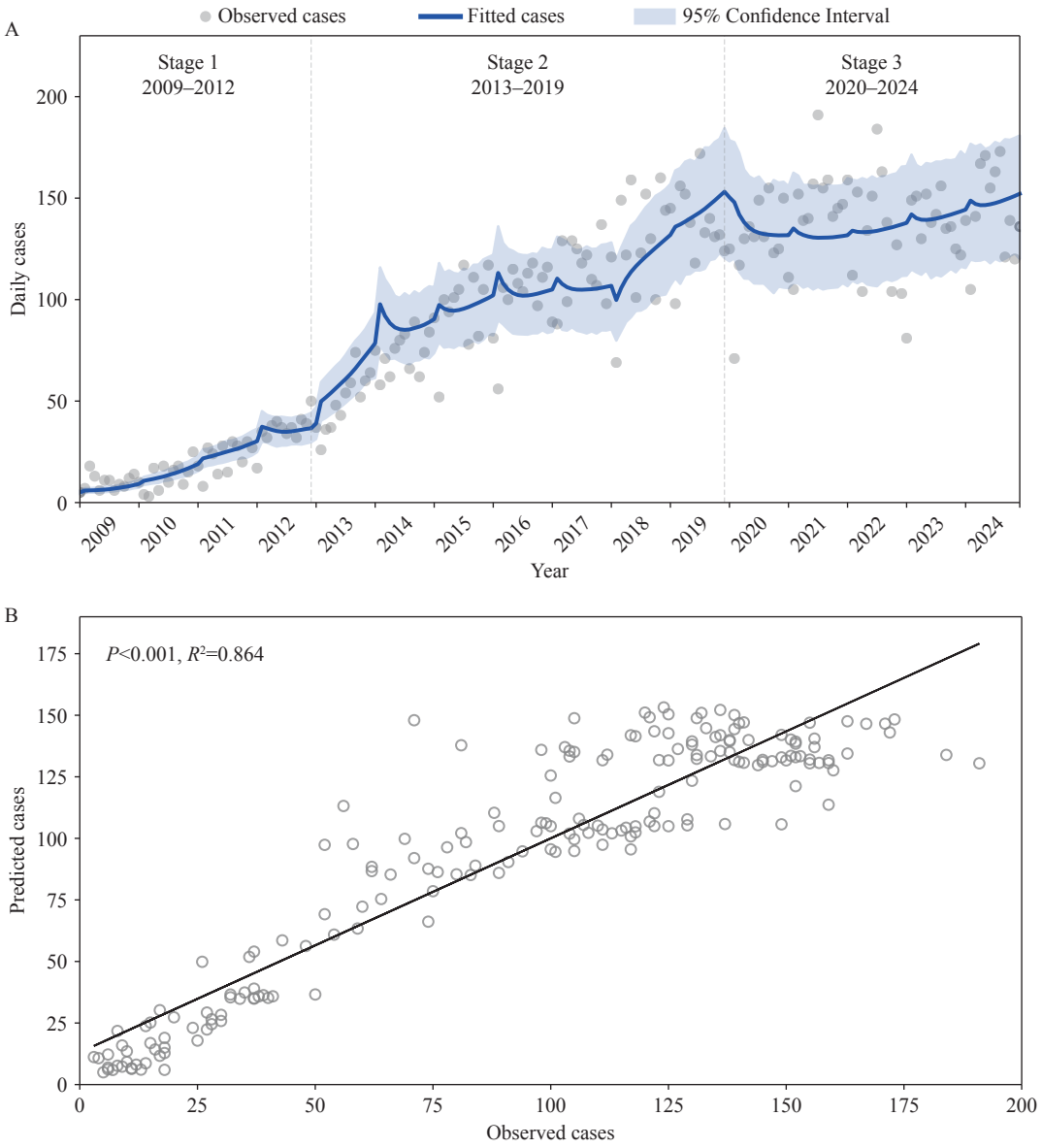
$$\frac{dT_1}{dt}=p_1I-p_2\delta T_1-(1-p_2)\omega_2T_1-(d_r+f_1)T_1$$

$$\frac{dT_2}{dt}=p_2\delta T_1-p_3\delta T_2(1-p_3)\omega_3T_2-(d_r+f_1)T_2$$

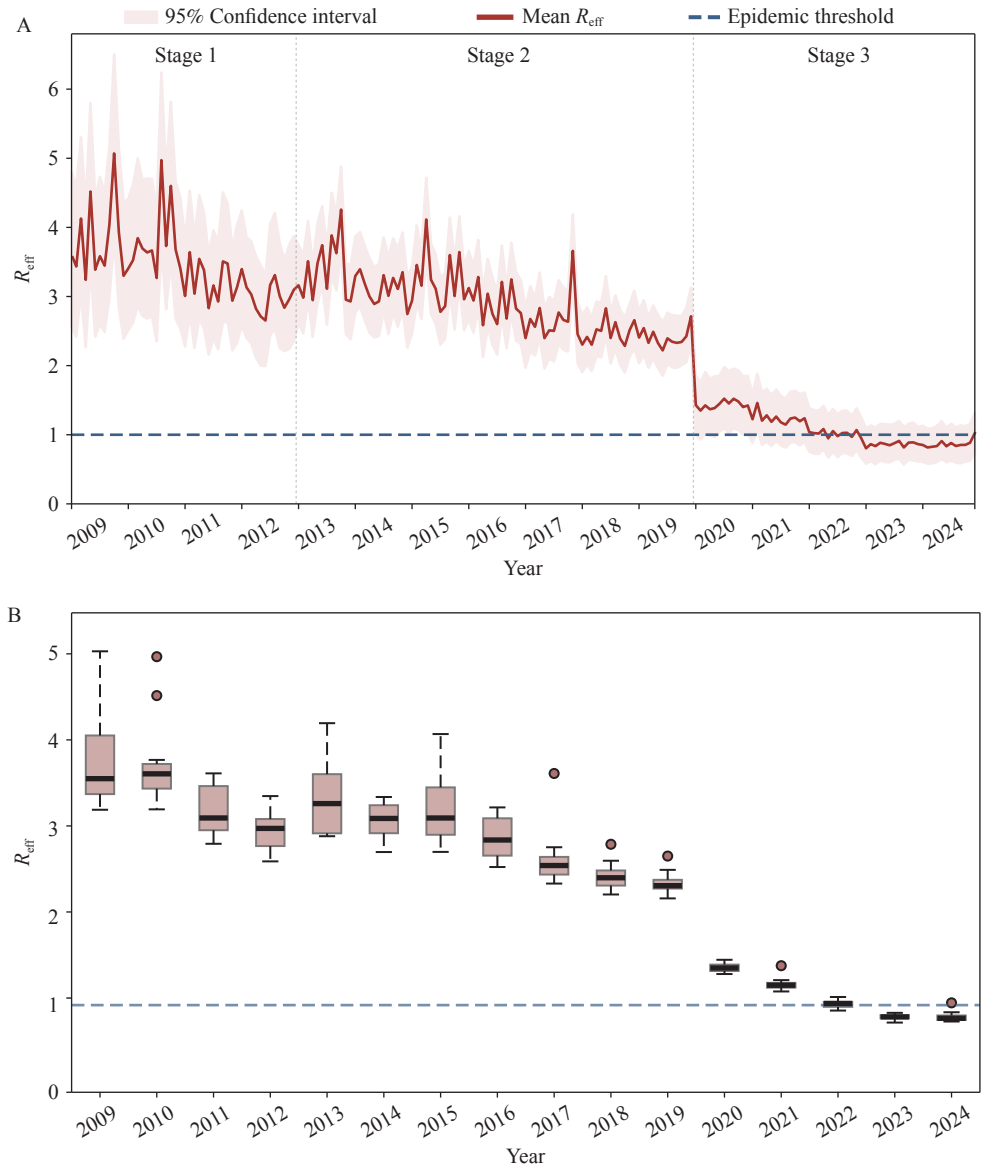
$$\frac{dT_3}{dt}=p_3\gamma T_2-q\omega_4T_3-d_rT_3$$

$$\frac{dA}{dt}=(1-p_1)\omega_1I+(1-p_2)\omega_2T_1+(1-p_3)\omega_3T_2+q\omega_4T_3-(d_r+f_2)A$$

SUPPLEMENTARY FIGURE S1. Structure of the HIV transmission dynamics model. (A) Schematic diagram of the SITD transmission dynamics model; (B) Differential equations of the SITD transmission dynamics model.
Abbreviation: HIV=human immunodeficiency virus.



SUPPLEMENTARY FIGURE S2. Model fitting results of the HIV transmission dynamics model in Fujian Province, 2009–2024. (A) Fitted cases with 95% confidence intervals and observed cases; (B) Correlation between predicted and observed cases (P<0.001, R^2=0.864).
Abbreviation: HIV=human immunodeficiency virus.



SUPPLEMENTARY FIGURE S3. HIV transmissibility in Fujian Province. (A) Monthly mean R_{eff} with 95% confidence intervals; (B) Annual mean R_{eff} distributions (boxplots). The dashed line represents the epidemic threshold (R_{eff}=1).
Abbreviation: HIV=human immunodeficiency virus.

Preplanned Studies

Time to Antiretroviral Therapy Initiation After HIV Diagnosis in the Integrase Strand Transfer Inhibitor Era — Nanjing City, Jiangsu Province, China, 2021–2024

Chenyu Ma[1]; Jinjin Yang[1]; Anni Liu[2]; Zheng Qian[2]; Ziyao Liu[1]; You Ge[2]; Yuanyuan Xu[3]; Zhengping Zhu[3]; Hongxia Wei[2]; Zhihang Peng[4,#]; Guoping Yin[1,#]; Zhiliang Hu[2,5,#]

Summary

What is already known about this topic?

In China, patients with human immunodeficiency virus are diagnosed at the CDC and then referred to hospitals for antiretroviral therapy (ART), but treatment delays persist. Previous studies have reported overall delay in time from diagnosis to ART initiation without identifying where delays occur.

What is added by this report?

A Nanjing cohort study showed that hospital delays were almost eliminated with integrase strand transfer inhibitor (INSTI)-based regimens: 87.9% of patients initiated ART within 7 days of their first hospital visit. However, the main bottleneck remains CDC referral, as 49.8% of patients attended the hospital within 7 days of HIV confirmation.

What are the implications for public health practice?

Shortening the CDC-to-hospital interval via stronger linkage and equitable INSTI access is critical for achieving rapid ART.

ABSTRACT

Introduction: In China, the total time from laboratory confirmation of human immunodeficiency virus (HIV) infection at the Center for Disease Control and Prevention to antiretroviral therapy (ART) initiation (T_total) comprises two intervals: the time from confirmed HIV diagnosis to the first hospital visit (T_PreHOS) and the time from the first hospital visit to ART initiation (T_PostHOS). However, the relative contributions of these intervals and their changes in the integrase strand transfer inhibitors (INSTI) era remain unclear.

Methods: This retrospective cohort study included adults with newly diagnosed HIV infection in Nanjing, China (2021–2024). Temporal trends in T_total, T_PreHOS, and T_PostHOS were assessed across calendar years. Factors associated with delays were examined using Cox proportional hazards model.

Results: Among the 1,456 patients analyzed, all three intervals showed substantial year-to-year reductions. The median T_total decreased from 20 to 9 days, driven by reductions in T_PreHOS (12 to 7 days) and T_PostHOS (7 to 0 days). In 2024, 37% initiated ART within 7 days of diagnosis, 49.8% attended the hospital within 7 days of HIV confirmation, and 87.9% initiated ART within 7 days of their first hospital visit. Notably, 56.3% started ART at their initial visit. Younger age was associated with delayed hospital attendance. INSTI-based regimens were independently associated with faster ART initiation.

Conclusion: Hospital-level delays have been largely eliminated in the INSTI era, making same-day ART initiation increasingly achievable. Further reductions in delay mandates targeted efforts to shorten T_PreHOS.

Rapid or same-day antiretroviral therapy (ART) initiation reduces viremia duration, disease progression, mortality, and human immunodeficiency virus (HIV) transmission (*1*). In China, patients are typically diagnosed at the Center for Disease Control and Prevention (CDC) and then referred to designated hospitals for ART; this diagnosis–referral–first-visit pathway may prolong time to ART initiation (*2*). New-generation integrase strand transfer inhibitors (INSTIs), with high resistance barriers and reduced reliance on baseline laboratory results, facilitate rapid ART after hospital presentation (*3*); however, real-world evidence on how the diagnosis-to-treatment interval (T_total) has changed in the INSTI era is limited. The relative contributions of the pre-hospital interval (from diagnosis to the first hospital visit;

T_PreHOS) and post-hospital interval (from the first hospital visit to ART initiation; T_PostHOS) remain unclear. Although INSTIs were recommended by WHO in 2015 and China's national guidelines in 2018 (*4*–*5*), uptake was constrained by cost until price reductions and reimbursement expansion in 2020 (*6*). We therefore define 2021–2024 as the "integrase strand transfer inhibitor (INSTI) era."

To address this gap, we conducted a retrospective cohort study of adults (≥18 years) with HIV infection diagnosed between 2021 and 2024 and managed at the Second Hospital of Nanjing, which provides care for approximately 90% of persons with HIV in Nanjing (local CDC data), excluding those with severe AIDS-defining illnesses before ART initiation. In routine outpatient care, the only prerequisite for ART initiation was the absence of major abnormalities in routine blood tests and basic chemistry results, indicating intolerance to potential drug-related adverse events; and no regimen-specific or protocolized rapid-ART pathway was implemented. Temporal trends in T_total, T_PreHOS, and T_PostHOS were analyzed and their relative contributions to the overall delay were assessed. Group differences were compared using the Wilcoxon rank-sum test or the Chi-square test, as appropriate. Temporal trends were assessed using the Jonckheere–Terpstra trend test.

To identify factors associated with longer or shorter intervals, time-to-event analyses were conducted with T_PreHOS and T_PostHOS, treated as continuous outcomes without pre-specified delay thresholds. Kaplan–Meier methods and Cox proportional hazards models were used to evaluate the time to first hospital visit and time to ART initiation. For T_PreHOS, time zero was the date of laboratory-confirmed HIV diagnosis, and the event was the first hospital visit; for T_PostHOS, time zero was the first hospital visit, and the event was ART initiation. Participants were followed up until event occurrence or administratively censored at 30 days (T_PreHOS) or 14 days (T_PostHOS), whichever occurred first. All analyses were conducted using the R software (version 4.4.2; R Foundation for Statistical Computing, Vienna, Austria).

During 2021–2024, 1,818 adults newly initiated ART; after excluding 362, 1,456 were included, of whom 1,369 (94.0%) were male, with a median age of 30 years [interquartile range (IQR): 24.7–40.4]. Most infections occurred via homosexual transmission (1,056, 72.5%). At baseline, the median plasma HIV RNA level was 30,050 copies/mL (IQR: 6,500–83,800), and the median CD4 count was 351 cells/μL (IQR: 232–475). T_total had a median duration of 15 days (IQR: 9–27). In total, 230 patients (15.8%) initiated ART within 7 days after HIV confirmation and 1,120 (76.9%) started treatment within 30 days. The median T_PreHOS was 10 days (IQR: 5–19); 516 (35.4%) and 1,224 (84.1%) visited a designated hospital within 7 days and 30 days after HIV confirmation, respectively. T_PostHOS was shorter, with a median of 5 days (IQR: 0–7); 402 patients (27.6%) initiated ART on the same day as their first hospital visit, 937 (64.4%) within 7 days, and 1,402 (96.3%) within 30 days (Table 1).

We then examined the temporal changes in these intervals across calendar years. Median T_total declined steadily from 20 days (IQR: 13–34) in 2021 to 18 (12–32) in 2022, 14 (7–26) in 2023, and 9 (5–18) in 2024. Median T_PreHOS is 12 days (IQR: 6–25) in 2021 and remain 12 (6–20) in 2022, followed by a decrease to 10 (5–19) in 2023 and 7 (4–13) in 2024. Median T_PostHOS decreased from 7 days (IQR: 4–8) in 2021 and 6 (3–8) in 2022 to 2 (0–6) in 2023 and 0 (0–3) in 2024 (Figure 1A–C).

Correspondingly, ART initiation within 7 days of diagnosis increased from approximately 7.0% in 2021 to 37.0% in 2024, whereas initiation within 30 days increased from 71.9% to 87.1% (Figure 1D). For T_PreHOS, the proportion of patients attending a designated hospital within 7 days increased from 30.9% to 49.8% and within 30 days from 80.0% to 91.0% (Figure 1E). For T_PostHOS, same-day initiation increased from 9.3% to 56.3%, and initiation within 7 days of the first visit increased from 45.8% to 87.9% (Figure 1F).

To assess stage contributions, we compared absolute durations and proportional contributions (T_PreHOS/T_total and T_PostHOS/T_total). Across all years, median T_PreHOS consistently exceeded T_PostHOS (all $P<0.05$). Accordingly, median T_PreHOS/T_total ratio increased from 0.65 (IQR: 0.46–0.86) in 2021 and 0.69 (0.50–0.92) in 2022 to 0.86 (0.60–1.00) in 2023 and 1.00 (0.70–1.00) in 2024, consistently exceeding the corresponding T_PostHOS/T_total ratios of 0.35 (IQR: 0.14–0.54), 0.31 (IQR: 0.08–0.50), 0.14 (IQR: 0–0.40), and 0 (IQR: 0–0.30), respectively (all $P<0.05$; Figure 1G). These findings indicate that pre-hospital delays accounted for most of the diagnosis-to-treatment interval.

For T_PreHOS, Kaplan–Meier curves showed significant variations in hospital attendance within 30

TABLE 1. Characteristics of the participants.

Characteristics	Total (*N*=1,456), *n* (%)
Age, years	30 (24.7–40.4)*
<30	730 (50.1)
30–44	446 (30.6)
45–59	174 (12.0)
≥60	106 (7.3)
Men	1,369 (94.0)
Marital status	
Unmarried	986 (67.7)
Married or living with a partner	364 (25.0)
Widowed, divorced, or unknown	106 (7.3)
Routes of HIV transmission	
HET	386 (26.5)
MSM	1,056 (72.5)
IDU or unknown	14 (1.0)
BMI, kg/m^2	22.4 (20.3–24.5)*
Occupation	
Employee outside the public-sector system	646 (44.4)
Informal or flexible employment	434 (29.8)
Public sector employee	197 (13.5)
Student	179 (12.3)
Educational attainment	
Junior college and above	946 (65.0)
High school and secondary vocational school	201 (13.8)
Junior high school and below	251 (17.2)
Unknown	58 (4.0)
VL, copies/mL	30,050 (6,500–83,800)*
<10,000	434 (29.8)
10,000–50,000	494 (33.9)
>50,000	528 (36.3)
CD4, cells/μL	351 (232–475)*
<350	721 (49.5)
≥350	735 (50.5)
T_PreHOS, days	10 (5.0–19.0)*
≤7	516 (35.4)
≤30	1,224 (84.1)
T_PostHOS, days	5 (0.0–7.0)*
same-day	402 (27.6)
≤7	937 (64.4)
≤30	1,402 (96.3)
T_total, days	15 (9.0–27.0)*
≤7	230 (15.8)
≤30	1,120 (76.9)
Initial ART regimen	
NNRTI based	472 (32.4)
INSTI-2DR	419 (28.8)
INSTI-3DR	557 (38.3)
Other	8 (0.5)

Note: * Data are presented as median (interquartile range).
Abbreviation: ART=antiretroviral therapy; HET=heterosexual contact; MSM=men who have sex with men; IDU=intravenous drug use; BMI=body mass index; VL=viral load; CD4=CD4 count; NNRTI=non-nucleoside reverse transcriptase inhibitor; INSTI-2DR=integrase strand transfer inhibitor–based two-drug regimen; INSTI-3DR=integrase strand transfer inhibitor–based three-drug regimen; T_PreHOS=time from confirmed HIV diagnosis to the first hospital visit; T_PostHOS=time from the first hospital visit to ART initiation; T_total=time from confirmed HIV diagnosis to ART initiation (T_total=T_PreHOS+T_PostHOS).

days of diagnosis according to age and marital status (all P<0.05; Supplementary Figure S1A–C, available at https://weekly.chinacdc.cn/). In the multivariable Cox regression analysis, younger age (<30 years) was independently associated with a lower risk of hospital attendance (Supplementary Table S1, available at https://weekly.chinacdc.cn/).

INSTI-based regimens were predominant (67.0%), followed by non-nucleoside reverse transcriptase inhibitor (NNRTI)-based regimens (32.4%). The proportion of patients receiving INSTI-based regimens increased from 52.6% in 2021 to approximately 73% in 2022 and remained stable thereafter, at approximately 75% at first hospital visit (Figure 2). For T_PostHOS, Kaplan–Meier analyses demonstrated differences in ART initiation speed across baseline CD4 strata and initial regimen types (all P<0.05; Supplementary Figure S1D–I, available at https://weekly.chinacdc.cn/). In the multivariable Cox model, baseline CD4 count and regimen remained significant: patients with CD4 ≥350 cells/μL had a lower hazard of ART initiation, whereas those receiving INSTI-based regimens had a higher hazard of ART initiation (Supplementary Table S2, available at https://weekly.chinacdc.cn/).

DISCUSSION

From 2021 to 2024, both T_PreHOS and T_PostHOS declined substantially, paralleling rapid ART implementation and expanded access to INSTI-based regimens after the 2021 insurance negotiations (*6*). Median T_PostHOS decreased from 7 days to 0

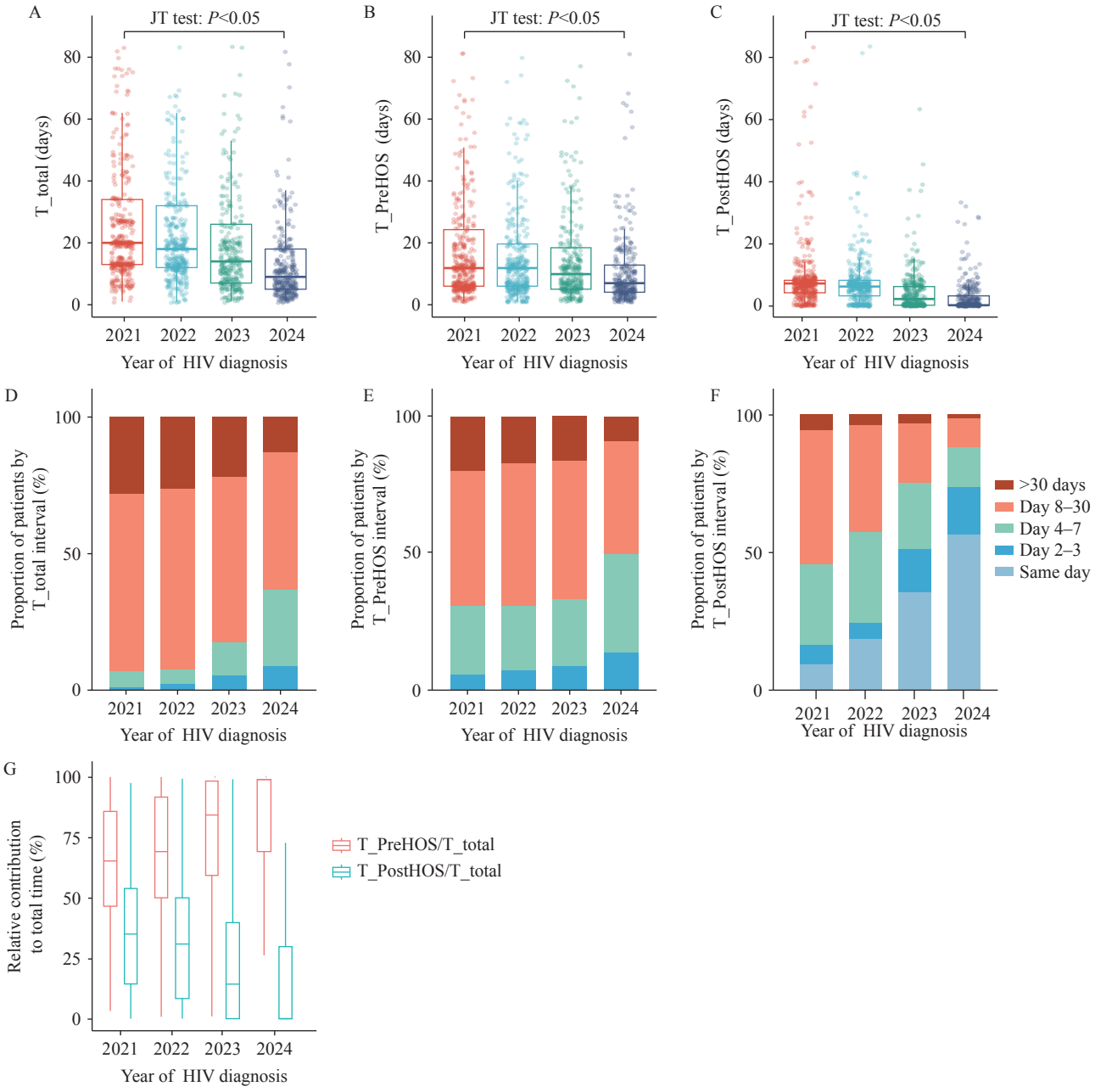


FIGURE 1. Time intervals from HIV diagnosis to ART initiation: annual trends and distributions among adults in nanjing, 2021–2024. (A–C) Boxplots of T_total, T_PreHOS, and T_PostHOS by diagnosis year. (D–F) Proportions of patients in each interval category for T_total, T_PreHOS, and T_PostHOS. (G) Relative contributions of T_PreHOS and T_PostHOS to T_total.
Note: T_total = T_PreHOS+T_PostHOS.
Abbreviation: ART=antiretroviral therapy; T_PreHOS=time from confirmed HIV diagnosis to the first hospital visit; T_PostHOS=time from the first hospital visit to ART initiation; T_total=time from confirmed HIV diagnosis to ART initiation; JT=Jonckheere-Terpstra.

days, with over half of the patients initiating ART on the day of their first visit since 2022, indicating that it was no longer a major barrier. In contrast, despite improvement, T_PreHOS remained prolonged, with more than half of the patients in 2024 not accessing care within 7 days of diagnosis, making it the main limiting factor for further reductions in T_total.

T_PreHOS reflected a predominantly patient-dependent process. In multivariable analysis, younger age (<30 years) was independently associated with a longer T_PreHOS. Delayed hospital presentation among younger patients may reflect lower perceived

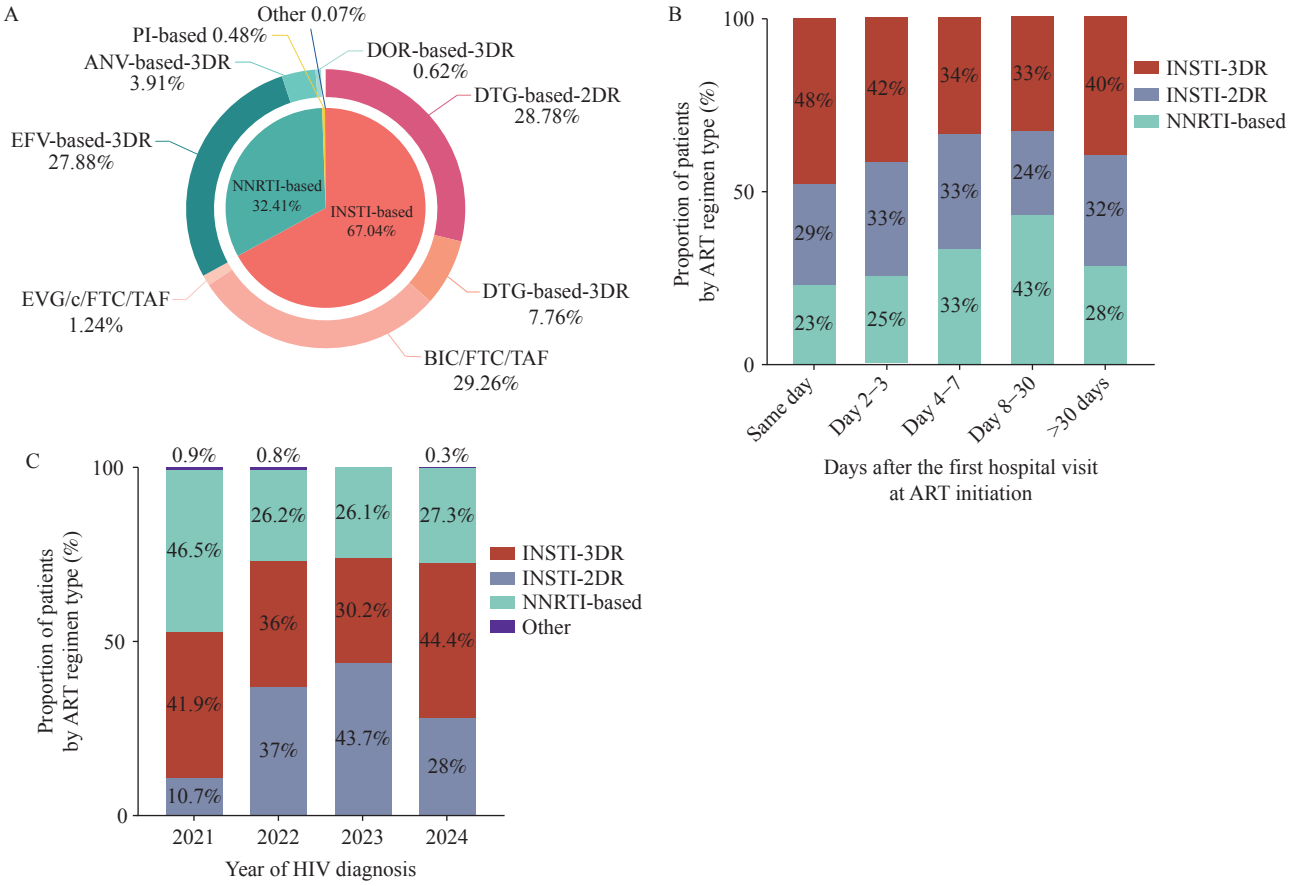


FIGURE 2. Distribution and temporal patterns of initial ART regimens. (A) Overall distribution of initial ART regimens. (B) Distribution of initial ART regimens by time from first hospital visit to ART initiation. (C) Distribution of initial ART regimens by year of HIV diagnosis.
Abbreviation: ART=antiretroviral therapy; BIC=bictegravir; DOR=doravirine; DTG=dolutegravir; EFV=efavirenz; EVG/c=elvitegravir/cobicistat; FTC=emtricitabine; TAF=tenofovir alafenamide; 3TC=lamivudine; ANV=ainuovirine; INSTI=integrase strand transfer inhibitor; INSTI-2DR=INSTI-based two-drug regimen; INSTI-3DR=INSTI-based three-drug regimen; NNRTI=non-nucleoside reverse transcriptase inhibitor; PI=protease inhibitor.

urgency, fewer symptoms, and psychological factors, such as denial or stigma. Targeted post-diagnosis counseling, youth-focused education, and streamlined referral navigation may help reduce pre-hospital delays.

In contrast, T_PostHOS decreased substantially over time and contributed minimally to T_total. After adjustment, the INSTI-based regimens were the main regimen-level factors associated with faster ART initiation. Rather than reflecting a predefined rapid ART program, this association likely reflects evolving physician-driven practice patterns in a resource-rich setting, where early initiation is increasingly favored after safety assessment. INSTI-based regimens are therefore more frequently selected for earlier initiation because of perceived safety and flexibility, without implying a mandatory or exclusive role in accelerating treatment.

For patients in whom INSTI-based regimens were not selected, regardless of patient preference or other practical considerations, ART initiation was not subject to mandatory delays or required laboratory results. However, both clinicians and patients may have opted for a less urgent initiation timeline, allowing additional clinical information to guide treatment decisions. Additionally, patients with higher baseline CD4 counts experienced modest delays in ART initiation, possibly reflecting a lower perceived urgency for immediate treatment after clinical assessment. Importantly, this study does not assess rapid ART intervention, but describes organically accelerated ART initiation under preserved clinical discretion in a resource-rich environment.

This observed delay refines prior findings across different settings. National surveillance data from China show substantial reductions in the overall diagnosis-to-treatment interval, but aggregate pre- and

post-hospital phases, obscuring phase-specific determinants (*7*). Building on the Guangxi "simplified test and treat" intervention, which identified service fragmentation as a key driver of delay and mortality, our analysis indicates that administrative separation between CDC confirmation and hospital-based treatment remains the principal contemporary bottleneck (*8*). Similar structural barriers have been reported in other systems, although with distinct system architectures; streamlined pathways in Hong Kong SAR, and rapid-entry programs in the United States have shortened diagnosis-to-ART intervals by reducing administrative complexity (*9–10*). Within China's referral framework, integrated public health–clinical models could further shorten T_PreHOS including on-site joint CDC–hospital services enabling confirmatory testing, referral, and ART assessment within a single clinical unit, as well as electronic referral and shared data platforms allowing real-time appointment scheduling at diagnosis (*8*).

The findings in this report are subject to at least three limitations. First, this was a single-center study that primarily included clinically stable, outpatient adults initiating ART; therefore, the findings may not be generalizable to individuals with advanced disease or inpatient care pathways. Second, several unmeasured patient-level and system-related determinants could not be explicitly modeled, impeding quantification of their effects on T_PreHOS. Finally, the favorable local infrastructure may further limit direct generalizability to resource-limited settings, although, expanding access to generic INSTI-based regimens, may enable similar trends elsewhere.

In summary, in the INSTI era, China's ART strategy has progressed, but T_PreHOS remains limiting; achieving same-day or 7-day ART initiation requires integrated public health and clinical services, institutions and patient engagement to reach national 95–95–95 targets.

Ethical statement: Approved by the Ethics Committee of the Second Hospital of Nanjing (2024-LS-ky076). All the participants provided written informed consent.

Conflict of interest: No conflicts of interest.

Funding: Supported by Jiangsu Provincial Health Commission Medical Research Foundation (M2022030), and the General Clinical Trial Project of Nanjing Medical Science and Technology Development Fund (LYM25017).

doi: 10.46234/ccdcw2026.078

[#] Corresponding authors: Zhiliang Hu, huzhiliang@njucm.edu.cn; Guoping Yin, yinguoping0304@163.com; Zhihang Peng, pengzh@chinacdc.cn.

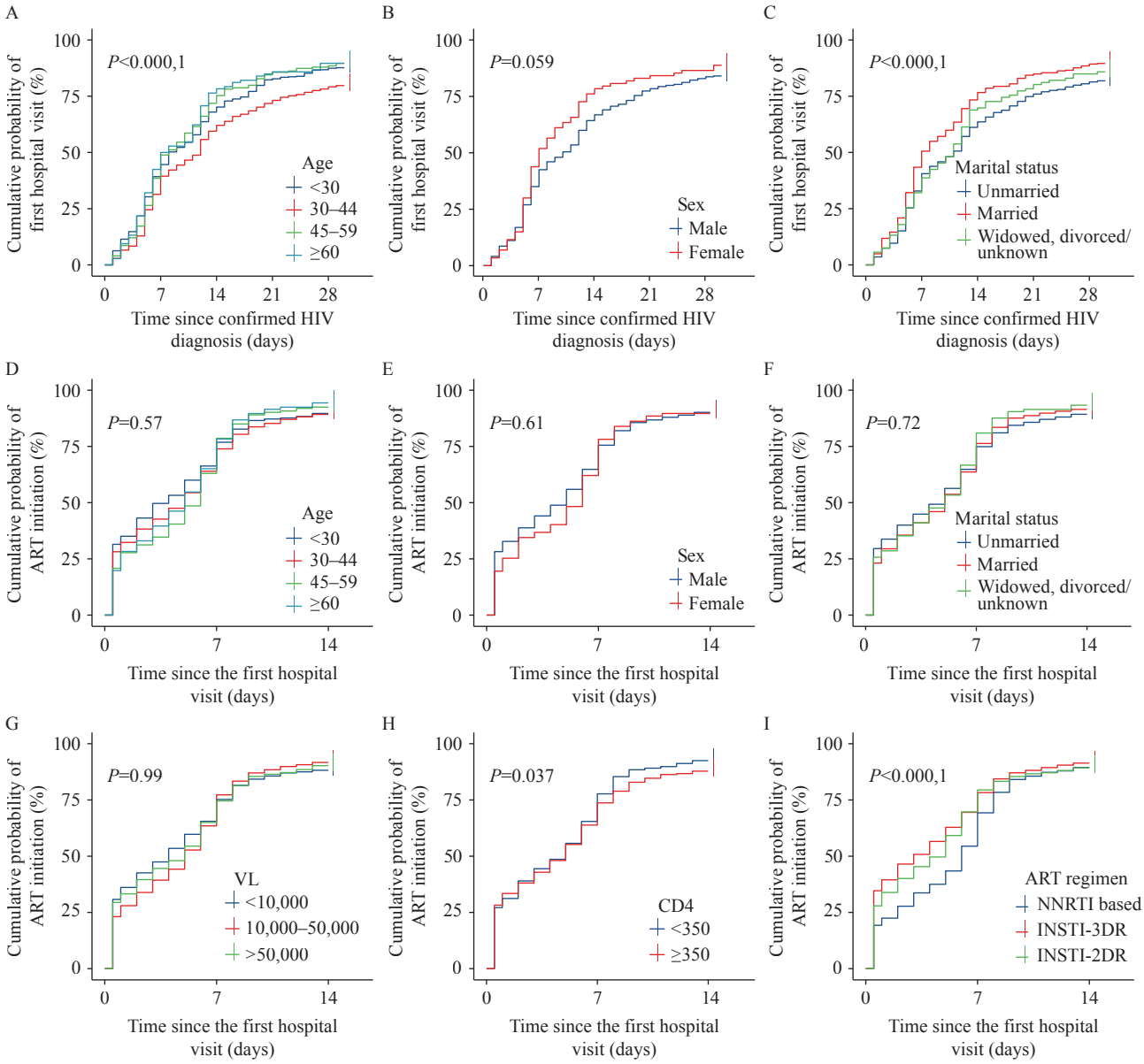
[1] School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China.; [2] Department of Infectious Diseases, the Second Hospital of Nanjing, Nanjing University of Chinese Medicine, Nanjing City, Jiangsu Province, China.; [3] Department of AIDS/STD Control and Prevention, Nanjing Municipal Center for Disease Control and Prevention, Nanjing City, Jiangsu Province, China; [4] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China; [5] Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China.

Submitted: November 26, 2025
Accepted: March 04, 2026
Issued: April 17, 2026

REFERENCES

1. Kamya MR, Petersen ML, Kabami J, Ayieko J, Kwariisima D, Sang N, et al. SEARCH human immunodeficiency virus (HIV) streamlined treatment intervention reduces mortality at a population level in men with low CD4 counts. Clin Infect Dis 2021;73(7):e1938 – 45. https://doi.org/10.1093/cid/ciaa1782.
2. Vu TT, Rupasinghe D, Khol V, Chaiwarith R, Tanuma J, Kumarasamy N, et al. Temporal trends from HIV diagnosis to ART initiation among adults living with HIV in the Asia-Pacific (2013-2023). AIDS Res The 2025;22(1):29. https://doi.org/10.1186/s12981-025-00718-8.
3. McKellar MS, Keys JR, Filiatreau LM, McGee KS, Kuruc JD, Ferrari G, et al. Rapid viral suppression using integrase inhibitors during acute HIV-1 infection. J Antimicrob Chemother 2025;80(1):169 – 74. https://doi.org/10.1093/jac/dkae391.
4. World Health Organization. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. Geneva: World Health Organization; 2015.
5. AIDS and Hepatitis C Professional Group, Society of Infectious Diseases, Chinese Medical Association, Chinese Center for Disease Control and Prevention. Chinese guidelines for diagnosis and treatment of HIV/AIDS (2018). Chin J Intern Med 2018;57(12):867-84. http://dx.doi.org/10.3760/cma.j.issn.0578-1426.2018.12.002. (In Chinese).
6. Zhu Z, Zhang JW, Xu ZH, Wang Q, Qi Y, Yang L. Impacts of National Reimbursement Drug Price Negotiation on drug accessibility, utilization, and cost in China: a systematic review. Int J Equity Health 2025;24(1):36. https://doi.org/10.1186/s12939-025-02390-w.
7. Zhang JJ, Zhou MG, Wang PC, Wang DD, Mi YQ, Liu JF, et al. Temporal trends of clinical characteristics and treatments in people living with HIV at the initiation of antiretroviral therapy - Beijing municipality, China, 2010-2020. China CDC Wkly 2023;5(6):131 – 6. https://doi.org/10.46234/ccdcw2023.024.
8. Wu ZY, Zhao Y, Ge XM, Mao YR, Tang ZZ, Shi CX, et al. Simplified HIV testing and treatment in China: analysis of mortality rates before and after a structural intervention. PLoS Med 2015;12(9):e1001874. https://doi.org/10.1371/journal.pmed.1001874.
9. Poon PKM, Wong NS, Leung WS, Wong BCK, Kwong TS, Kwan TH, et al. The differential impacts of early detection and accelerated antiretroviral therapy on the epidemiologic trend of sexually acquired HIV infection in Hong Kong. PloS One 2022;17(9):e0274498. https://doi.org/10.1371/journal.pone.0274498.
10. Colasanti J, Sumitani J, Mehta CC, Zhang YR, Nguyen ML, Del Rio C, et al. Implementation of a rapid entry program decreases time to viral suppression among vulnerable persons living with HIV in the southern United States. Open Forum Infect Dis 2018;5(6):ofy104. https://doi.org/10.1093/ofid/ofy104.

SUPPLEMENTARY MATERIAL



SUPPLEMENTARY FIGURE S1. Cumulative probability of first hospital visit and ART initiation following HIV diagnosis. Kaplan-Meier curves show the cumulative probability of (A–C) the first hospital visit after laboratory-confirmed HIV diagnosis (T_PreHOS) and (D–I) ART initiation after the first hospital visit (T_PostHOS), stratified by demographic and clinical characteristics.

Note: Panels A–C present time from confirmed HIV diagnosis to the first hospital visit, stratified by age (A), sex (B) and marital status (C). Panels D-I present time from the first hospital visit to ART initiation, stratified by age (D), sex (E), marital status (F), baseline viral load (G), baseline CD4 cell count (H), and initial ART regimen type (I). Participants were followed for up to 30 days after HIV diagnosis for T_PreHOS and up to 14 days after the first hospital visit for T_PostHOS. Individuals who did not experience the corresponding event within the predefined observation window were right-censored. *P* values were calculated using the log-rank test.

Abbreviation: ART=antiretroviral therapy; VL=viral load; CD4=CD4 count; INSTI=integrase strand transfer inhibitor; INSTI-2DR=INSTI-based two-drug regimen; INSTI-3DR=INSTI-based three-drug regimen; NNRTI=non-nucleoside reverse transcriptase inhibitor.

SUPPLEMENTARY TABLE S1. Factors associated with time to first hospital visit after HIV diagnosis (T_PreHOS) among adults initiating ART in Nanjing, China, 2021–2024.

Variable	No. of events/patients	Univariable		Multivariable	
		HR (95% *CI*)	*P*	aHR (95% *CI*)	*P*
Age, years					
30–44	391/446	Ref		Ref	
<30	582/730	0.77 (0.67, 0.87)	<0.001	0.79 (0.68, 0.92)	0.002
45–59	156/174	1.06 (0.88, 1.28)	0.525	0.94 (0.78, 1.20)	0.789
≥60	95/106	1.11 (0.89, 1.39)	0.348	1.02 (0.79, 1.33)	0.857
Sex					
Male	1,147/1,369	Ref		Ref	
Female	77/87	1.25 (1.00, 1.58)	0.055	1.11 (0.87, 1.43)	0.410
Marital status					
Unmarried	807/986	Ref		Ref	
Married or living with partner	326/364	1.36 (1.20, 1.55)	<0.001	1.14 (0.94, 1.37)	0.185
Widowed, divorced, or unknown	91/106	1.12 (0.90, 1.39)	0.314	0.92 (0.71, 1.19)	0.535
Routes of HIV transmission					
HET	328/386	Ref		–	–
MSM	885/1,056	0.96 (0.85, 1.09)	0.530	–	–
IDU or unknown	11/14	0.86 (0.47, 1.57)	0.619	–	–
Occupation					
Employee outside the public-sector system	546/646	Ref		Ref	
Public sector employee	170/197	1.17 (0.99, 1.40)	0.067	1.10 (0.92, 1.32)	0.303
Informal or flexible employment	369/434	1.06 (0.93,1.21)	0.411	1.02 (0.89,1.17)	0.768
Student	139/179	0.87 (0.73, 1.05)	0.156	1.01 (0.82, 1.23)	0.958
Educational attainment					
Junior college or above	798/946	Ref		–	–
High school or secondary vocational	167/201	0.94 (0.79, 1.11)	0.434	–	–
Middle school or below	211/251	1.01 (0.87, 1.17)	0.927	–	–
Unknown	48/58	1.12 (0.83, 1.49)	0.461	–	–

Note: "-" means not applicable. Hazard ratios (HRs) and adjusted hazard ratios (aHRs) with 95% confidence intervals were estimated using Cox proportional hazards models, with first hospital visit after laboratory-confirmed HIV diagnosis as the outcome, time zero defined as diagnosis, and administrative censoring at 30 days. All tests were two-sided (P<0.05).

Abbreviation: T_PreHOS=time from confirmed HIV diagnosis to the first hospital visit; ART=antiretroviral therapy; HIV=human immunodeficiency virus; HET=heterosexual contact; MSM=male-to-male sexual contact; IDU=injection drug use; Ref=reference category.

SUPPLEMENTARY TABLE S2. Factors associated with time to ART initiation after the first hospital visit (T_PostHOS) among adults initiating ART in Nanjing, China, 2021–2024.

Variable	No. of events/patients	Univariable		Multivariable	
		HR (95% *CI*)	*P*	aHR (95% *CI*)	*P*
Age, years					
30–44	399/445	Ref		–	–
<30	646/724	0.92 (0.81,1.05)	0.207	–	–
45–59	160/173	0.95 (0.79,1.14)	0.603	–	–
≥60	100/106	1.01 (0.81,1.26)	0.910	–	–
Sex					
Male	1,227/1,361	Ref		–	–
Female	78/87	0.94 (0.75,1.19)	0.623	–	–
Marital status					
Unmarried	875/980	Ref		–	–
Married or living with partner	332/363	1.01 (0.89,1.15)	0.851	–	–
Widowed, divorced, or unknown	98/105	1.09 (0.89,1.35)	0.396	–	–
Routes of HIV transmission					
HET	344/385	Ref		Ref	
MSM	948/1,049	1.11 (0.98,1.26)	0.090	1.11 (0.97, 1.27)	0.119
IDU or unknown	13/14	1.33 (0.76, 2.31)	0.317	1.33 (0.77, 2.32)	0.309
Occupation					
Employee outside the public-sector system	570/643	Ref		–	–
Public sector employee	182/196	1.08 (0.91, 1.28)	0.362	–	–
Informal or flexible employment	393/432	1.04 (0.91, 1.18)	0.576	–	–
Student	160/177	1.06 (0.89, 1.26)	0.534	–	–
Educational attainment					
Junior college or above	846/941	Ref		Ref	
High school or secondary vocational	181/199	0.89 (0.76, 1.05)	0.163	0.90 (0.76, 1.06)	0.200
Middle school or below	224/250	0.89 (0.77, 1.03)	0.110	0.93 (0.80, 1.10)	0.410
Unknown	54/58	1.37 (1.04, 1.80)	0.025	1.30 (0.98, 1.72)	0.066
Baseline VL, copies/mL					
<10,000	381/432	Ref		–	–
10,000–50,000	452/493	0.99 (0.87, 1.14)	0.926	–	–
>50,000	472/523	0.99 (0.87, 1.14)	0.936	–	–
Baseline CD4, cells/μL					
<350	664/718	Ref		Ref	
≥350	641/730	0.89 (0.80, 0.99)	0.035	0.87 (0.78, 0.98)	0.016
Regimen at ART initiation					
NNRTI based	422/472	Ref		Ref	
INSTI-3DR	509/557	1.33 (1.17, 1.51)	<0.001	1.30 (1.14, 1.48)	<0.001
INSTI-2DR	374/419	1.22 (1.06, 1.40)	0.006	1.20 (1.04, 1.39)	0.011

Note: "-" means not applicable. Hazard ratios (HRs) and adjusted hazard ratios (aHRs) with 95% confidence intervals were estimated using Cox models, with ART initiation after the first hospital visit as the outcome and administrative censoring at 14 days. Analyses of regimen type excluded 8 participants with non-classifiable regimens (analytic *n*=1,448). All tests were two-sided (*P*<0.05).

Abbreviation: ART=antiretroviral therapy; T_PostHOS=time from the first hospital visit to ART initiation; HIV=human immunodeficiency virus; HET=heterosexual contact; MSM=male-to-male sexual contact; IDU=injection drug use; VL=viral load; CD4=CD4 count; NNRTI=non-nucleoside reverse transcriptase inhibitor; INSTI-3DR=INSTI-based three-drug regimen; INSTI-2DR=INSTI-based two-drug regimen; Ref=reference category.

Methods and Applications

Development of a Subsequence Correlation Coefficient Feature Vector Method for High-Resolution HIV-1 Subtype Classification — China, 2004–2022

Shuyan Han[1,&]; Lily He[1,&]; Yihang Tang[2]; Kun Peter Li[3]; Yuhua Ruan[4]; Hengjian Cui[5,#]

Summary

What is already known about this topic?

Current HIV-1 subtype classification tools often rely on time-consuming alignment, whereas new non-alignment methods typically target single genes. China lacks a model for specifically predicting the non-B subtype strains prevalent domestically.

What is added by this report?

We developed a fast alignment-free method (SCCFV-RDA) for building multigene models. It achieved over 99.85% accuracy in classifying an international dataset and 99.7% accuracy in classifying Chinese *pol* gene data and showed superior recall for key circulating recombinant form subtypes.

What are the implications for public health practice?

This tool provides accurate and efficient computational support for the precise molecular surveillance of HIV-1 in China, thereby facilitating the formulation of targeted prevention and control strategies.

ABSTRACT

Introduction: The subtype classification of human immunodeficiency virus type 1 (HIV-1) is vital for its prevention and control. Current methods often rely on time-consuming sequence alignments, whereas new alignment-free approaches typically focus on single genes. The prevalent HIV strains in China are mainly non-B subtypes. However, no subtype prediction model exists for local sequences. Therefore, we aimed to develop a fast and accurate method for building multigene models specifically tailored to predict HIV-1 subtypes on the basis of Chinese data.

Methods: Herein, we propose a novel sequence-feature extraction method, named Subsequence Correlation Coefficient Feature Vector (SCCFV), which captures the spatial distribution and correlations of nucleotides and converts DNA sequences into high-dimensional numerical features. It then combines these features with a regularized discriminant analysis (RDA) classifier to build a dedicated classification model.

Results: The proposed SCCFV-RDA model exhibited robust and generalizable capabilities, maintaining more than 99.85% accuracy across an international dataset covering 10 gene regions. It also achieved 99.7% classification accuracy on an independent test set of Chinese HIV-1 *pol* gene data, showing significantly higher recall for mainstream circulating recombinant form subtypes than that achieved by traditional tools.

Conclusion: We developed a classification tool for HIV-1 subtypes and built a specialized model for HIV-1 strains prevalent in China. The accuracy and efficiency of the tool surpassed those of existing traditional methods, providing reliable computational support for the precise molecular epidemiological surveillance of HIV in China. This method holds significant practical value for facilitating the formulation of targeted prevention and control strategies.

According to recent statistics, approximately 40.8 million people worldwide were living with the human immunodeficiency virus (HIV) in 2025 (*1*). Increasingly complex and diverse HIV subtypes have emerged from the large-scale and evolving pandemic (*2*), posing numerous prevention and treatment challenges. As China is one of the key HIV-affected countries, the classification of its indigenous viral subtypes warrants attention. Through surveillance efforts from 2004 to 2022, the Chinese Center for Disease Control and Prevention (CDC) has established one of the largest national molecular epidemiology databases for acquired immunodeficiency syndrome (AIDS), currently providing data from 57,902 HIV-infected individuals (*3*). Domestically prevalent non-B

subtypes, such as CRF01_AE and CRF07_BC, exhibit significant genetic differences from the B subtype dominant in Europe and America (*4*). However, because the current mainstream classification tools are primarily trained on datasets dominated by European and American B subtypes and known circulating recombinant forms (CRFs), their performance in classifying the strains circulating in China may be limited, in particular for new CRFs that are not represented in their training data. Consequently, novel computational models that are specifically tailored to the genetic sequences prevalent in China and capable of efficiently processing large-scale data are urgently needed.

Machine learning and deep learning methods have been extensively applied to bioinformatics classification, demonstrating significant potential in this task (*5*). Existing models are predominantly trained on European and American B subtype data and typically target specific *pol* gene or full-genome regions, lacking a unified framework adaptable to diverse fragments. Herein, we propose a novel modular framework for HIV-1 subtype classification. Based on the Subsequence Correlation Coefficient Feature Vector (SCCFV) method, this framework establishes a model system adaptable to multiple gene fragments. To assess the applicability of the model beyond the *pol* gene region, we additionally collected 821 full-genome sequences from Chinese public databases for independent model validation. The model demonstrated outstanding performance in classifying complex CRFs in China, providing a unified and efficient solution for global HIV-1 genotyping and evolutionary surveillance based on arbitrary fragments.

METHODS

Dataset

The HIV-1 genome, which spans approximately 9,800 bp, comprises the highly conserved structural genes *gag*, *pol*, and *env* along with multiple regulatory and accessory genes. In this study, the sequence data for each gene segment were downloaded from the HIV Sequence Database, totaling 10 datasets. Categories with sample sizes exceeding 9 were selected to establish the models.

Of the 57,902 HIV *pol* gene sequences in the Chinese CDC AIDS database, 5% are labeled as "unique recombinant forms (URFs)," "Other," and "Other CRFs." Given their limited epidemiological significance, only sequences of the major HIV subtypes circulating in China were selected (Supplementary Figure S1, available at https://weekly.chinacdc.cn/).

Subsequence Correlation Coefficient Feature Vector Method

To analyze the base position distribution of the nucleotides in DNA sequences, we used a feature mapping method to transform the sequences into vector representations within a multidimensional feature space. For a given HIV DNA sequence, $G=(g_1,g_2,\cdots,g_n)$, where each nucleotide variable is $g_i = \{G, T, A, C\}$ and the sequence length is n. For a specific base, where $g \in \{A, T, C, G\}$, each base position is converted into a numerical feature:

$$v_g(i) = \begin{cases} 1, g_i = g \\ 0, g_i \neq g \end{cases} \quad i = 1,2,\ldots n. \tag{1}$$

Digital representation of the sequences is achieved by constructing four base-specific time series: $V_g = [V_g(1), V_g(2), \ldots, V_g(n)]$. For example, the sequence "AGCTAAG" can be converted as follows:

$$\begin{aligned} &V_A = [1, 0, 0, 0, 1, 1, 0] \\ &(\text{A appears in the 1st, 5th, and 6th positions.}) \end{aligned} \tag{2}$$

V_G, V_C, and V_T are similarly derived.

Two key statistical measures are calculated from the obtained numerical data: the average frequency and correlation coefficient. Let N_g denote the total count of nucleotide g in the sequence, where $g \in \{A, T, C, G\}$. The average frequency f_g represents the global proportion of base g, reflecting its compositional characteristics, and is defined as follows (Figure 1):

$$f_g = \frac{N_g}{n} \tag{3}$$

For a more in-depth analysis of the spatial distribution patterns of the bases, autocorrelation and cross-correlation functions were introduced. The autocorrelation function measures the strength of the association between the same bases after an interval of M positions:

$$\varnothing_{AA}(M) = \frac{1}{n}\sum_{i=1}^{n}(v_A(i) - f_A)(v_{A+M}(i) - f_A) \tag{4}$$

The autocorrelation coefficient $R_{AA}(M)$ of base A is defined as follows:

$$R_{AA}(M) = \frac{\varnothing_{AA}(M)}{\varnothing_{AA}(0)} \tag{5}$$

The M-step-delayed normalized cross-correlation coefficient between nucleotides A and T serves as a crucial metric for quantifying the strength of the

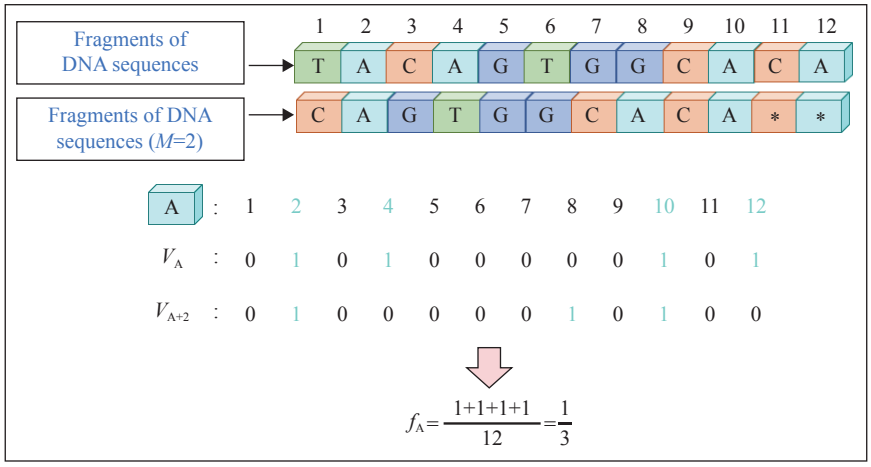


FIGURE 1. Computational methodology for determining V_A, V_{A+2}, and f_A.
Abbreviation: DNA=deoxyribonucleic acid.

correlation between two nucleotides at specific intervals within a sequence. The cross-correlation function is defined as follows:

$$\varnothing_{AT}(M) = \frac{1}{n}\sum_{i=1}^{n}(v_A(i) - f_A)(v_{T+M}(i) - f_T) \quad (6)$$

The correlation coefficient $R_{AT}(M)$ is defined as follows:

$$R_{AT}(M) = \frac{\varnothing_{AT}(M)}{\sqrt{\varnothing_{AA}(0) \cdot \varnothing_{TT}(0)}} \quad (7)$$

This coefficient reveals the periodicity of the base occurrence patterns; a high $R_{AA}(M)$ value indicates a tendency for the A bases to repeat at every M position. Cross-correlation quantifies the spatial dependencies between different bases, with $R_{AT}(M)$ reflecting the cooperative or antagonistic effects between A and T at M position intervals.

To comprehensively capture local feature variations across different regions of the DNA sequence, a segmented processing strategy is used. Specifically, each raw DNA sequence $G = (g_1, g_2, \cdots, g_n)$ is first uniformly divided into J contiguous subsequences. The first r subsequences (Substr$_1$, Substr$_2$, · · ·, Substr$_r$) each contains Y+1 nucleotides, whereas the remaining J-r subsequences (Substr$_{r+1}$, Substr$_{r+2}$, · · ·, Substr$J$) each contain Y nucleotides. The formula is as follows:

$$Y = \left[\frac{n}{J}\right], r = n - J \times Y (0 \le r \le J) \quad (8)$$

The standardized cross-correlation coefficient R for all possible nucleotide combinations (AA, AC, AG, AT, ..., TT; 16 combinations in total) is independently calculated for each subsequent segment at different M delay steps, resulting in a 16 × M-dimensional vector (*6*). The resulting eigenvalues are concatenated in segment order to form a high-dimensional feature vector with a total dimension of J × 16 × M.

Parameter Tuning and Classifier Selection

Four classifiers from machine learning methods (*7*) — Random Forest, XGBoost, Regularized Discriminant Analysis (RDA), and LightGBM — were selected for comparative analysis. To comprehensively validate the model performance, the following systematic evaluation framework was used:

$$\text{Accuracy} = \frac{|TP| + |TN|}{|TP| + |TN| + |FP| + |FN|} \quad (9)$$

$$\text{Precision} = \frac{|TP|}{|TP| + |FP|} \quad (10)$$

$$\text{Recall} = \frac{|TP|}{|TP| + |FN|} \quad (11)$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

As an example, we randomly divided 55,261 *pol* gene data points from the Chinese CDC database into 75% and 25% for use as training and independent test sets, respectively. Five-fold cross-validation was performed within the training set. This process simultaneously optimizes the hyperparameters of the classifier and the core parameters of the SCCFV feature extraction method. The classification performance of the final model was objectively evaluated using the independent test set, which was not involved in either the training or parameter selection processes.

According to the accuracy heatmaps of the four classifiers (Supplementary Figure S2, available at

https://weekly.chinacdc.cn/), RDA achieved the highest accuracy (99.7%) under the parameters J = 3 and M = 5 and was therefore selected as the final model with this parameter combination.

Figure 2 illustrates the overall framework of the SCCFV method. First, the input sequence is converted into a four-dimensional vector representation via a numerical mapping layer to capture base position information. Subsequently, the sequence is uniformly partitioned into J subsegments, and both local and global features are extracted via a statistical feature computation layer to construct a multidimensional feature vector. After input to the RDA classifier optimized via 5-fold cross-validation, the optimal model is obtained by parameter tuning, ultimately achieving precise HIV-1 subtype classification.

Comparison with Existing Tools

The performance of the SCCFV-RDA method was compared with those of COMET (*9*), REGA (*10*), and HIVdb (*11*), which are sequence alignment-based tools that are widely used as authoritative standards in the field. SNV (*12*), an alignment-free method based on single-nucleotide variation features, was also included for comparison. REGA was not included in the comparative analysis of the Chinese *pol* dataset because of its prohibitively long computation time.

Implementation Details

All the computational experiments were performed on a Lenovo Legion Y7000P laptop equipped with an Intel Core i7-14650HX processor, 16 GB RAM, and an NVIDIA GeForce RTX 4050 GPU. The algorithms were implemented using Python (version 3.9) in the Spyder integrated development environment. The key Python libraries NumPy and Pandas were used for data manipulation, whereas Scikit-learn (version 1.2; INRIA, Paris, France) was used for RDA, data preprocessing, model evaluation, and hyperparameter search (RandomizedSearchCV). The model persistence was assessed using Joblib. The source code and data are available at the repository provided in the "Data and code availability" section.

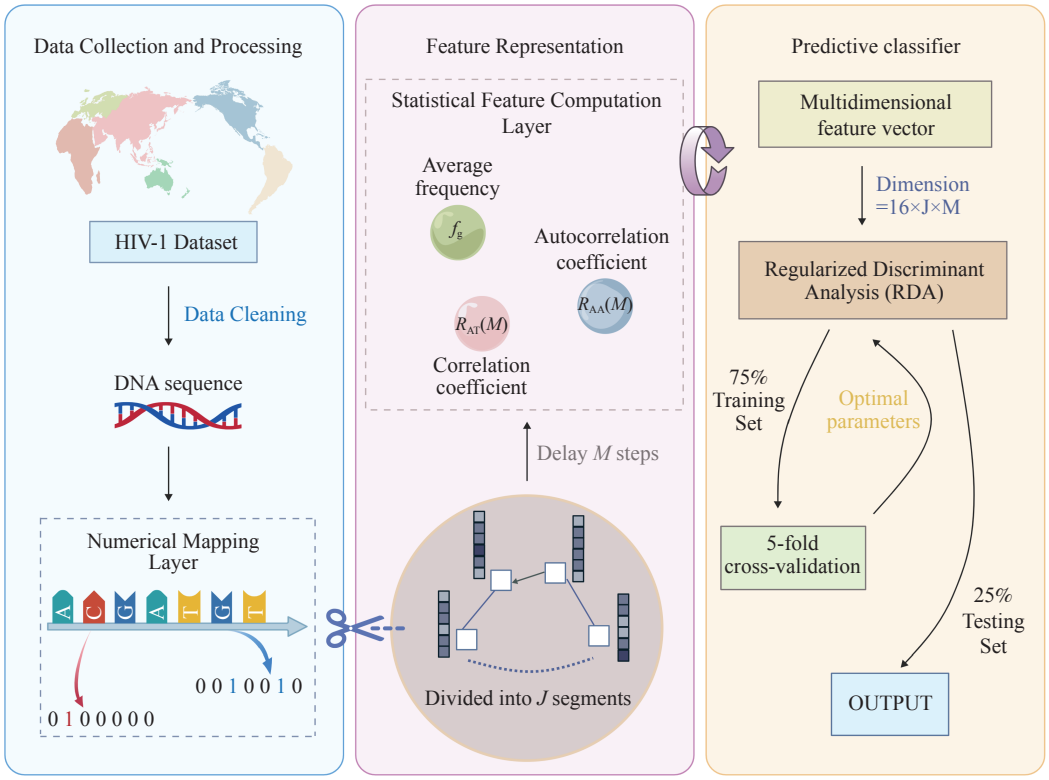


FIGURE 2. Framework diagram of the SCCFV method.
Note: created using the BioGDP tool (https://BioGDP.com) (8)
Abbreviation: HIV=human immunodeficiency virus; RDA=regularized discriminant analysis; SCCFV=subsequence correlation coefficient feature vector.

RESULTS

Classification of the HIV Sequence Datasets

We analyzed HIV subtypes with more than nine sequence entries for the full genome and *pol*, *env*, *gag*, *nef*, *rev*, *tat*, *vif*, *vpr*, and *vpu* genes. The model built for each gene demonstrated excellent and robust performance (Table 1).

Despite significant variations in the sample size and number of subtypes covered across the datasets, all gene segment models achieved accuracy rates exceeding 99.85%. This demonstrates that the proposed method possesses excellent generalization capability and robustness.

Classification of the Chinese Dataset

The model achieved 99.7% accuracy in classifying 25% of the Chinese HIV *pol* dataset used as an independent test set.

We compared our model with COMET, HIVdb, and SNV using this same *pol* independent test set (Table 2).

Compared with the HIVdb algorithm, which was developed using European and American B subtype data, our method demonstrated comparable performance in classifying the B subtype. However, for the CRF07_BC and CRF08_BC strains prevalent in China, our model achieved recall rates of 0.9882 and 0.9991, respectively, significantly outperforming HIVdb (recall: 0.6002 and 0.7873, respectively). Furthermore, compared with the COMET method that was based on European and American training data, our model clearly demonstrated better recall performance across all categories. The SNV-feature-based linear discriminant analysis classifier performed poorly, with the recall and F1-scores for all subtypes falling below those of the SCCFV method.

We further validated the SCCFV-RDA model using 821 full-genome sequences encompassing the major subtypes circulating in China. The model achieved 100% accuracy, recall, precision, and F1-score for each subtype, indicating classification performance (Supplementary Table S1, available at https://weekly.chinacdc.cn/). These results confirm that the SCCFV-RDA method generalizes effectively to non-*pol* genomic regions, reinforcing its utility for predicting HIV-1 strains prevalent in China.

TABLE 1. Classification metrics for 10 international datasets.

Dataset	ParameterJ	Parameter M	Accuracy	Recall	Precision	F1-score	Number of categories	Sample size
Complete	5	5	1	1	1	1	34	18,147
pol	5	5	0.9988	0.9988	0.9989	0.9988	35	30,708
env	5	5	0.9986	0.9986	0.9986	0.9986	15	159,312
gag	5	5	0.9997	0.9997	0.9997	0.9997	34	95,637
ref	5	4	0.9988	0.9988	0.9989	0.9988	19	56,826
rev	5	5	0.9985	0.9985	0.9985	0.9985	26	94,842
tat	4	5	1	1	1	1	32	41,346
vif	5	4	0.9987	0.9987	0.9988	0.9987	31	47,792
vpr	4	5	0.9993	0.9993	0.9993	0.9993	19	46,621
vpu	5	5	0.9997	0.9997	0.9997	0.9997	24	105,808

TABLE 2. Comparison of the performance of four models in classifying Chinese HIV *pol* data.

Method (Acc)	SCCFV (99.70%)			SNV (98.80%)			HIVdb (82.41%)			COMET (73.51%)		
Subtype	Recall	Precision	F1-score	Recall	Precision	F1-score	Recall	Precision	F1-score	Recall	Precision	F1-score
B	0.9920	0.9964	0.9942	0.9414	0.9833	0.9619	0.9973	0.9574	0.977	0.9574	1	0.9782
C	0.9892	0.9684	0.9787	0.8280	0.8105	0.8191	0.9892	0.7863	0.8762	0.7634	0.9861	0.8606
CRF01_AE	0.9989	0.9972	0.9981	0.9912	0.9862	0.9887	0.9870	0.9970	0.9920	0.7995	0.9996	0.8884
CRF07_BC	0.9982	0.9980	0.9981	0.9885	0.9818	0.9851	0.6002	0.9971	0.7494	0.5459	1	0.7063
CRF08_BC	0.9991	0.9973	0.9982	0.9873	0.9846	0.9859	0.7873	1	0.8810	0.9545	1	0.9767
CRF55_01B	0.9778	0.9925	0.9851	0.9275	0.9573	0.9421	0.8935	1	0.9438	0.892	1	0.9429

DISCUSSION

This is the first study to introduce a novel HIV-1 subtype classification method that integrates SCCFV with an RDA classifier. As an alignment-free approach, it effectively extracts discriminative features from HIV-1 sequences. It proved to be faster and more accurate than sequence alignment-based tools such as COMET and HIVdb and showed improvements across all classification metrics when evaluated against the alignment-free SNV method.

Our model — innovatively trained on a domestic sequence database — captures local genetic characteristics and addresses the limitations of international tools in classifying Chinese data. Moreover, unlike sequence alignment-based methods, the feature vector for each sequence need only be computed once, enabling the construction of a comprehensive feature vector database for all HIV-1 strains in China. Notably, the generalizability of our model beyond the *pol* region was confirmed using an independent set of 821 full-genome sequences.

Furthermore, the algorithm is designed to also predict subtypes for strains circulating globally. For instance, it achieved 99.88% classification accuracy on a global HIV-1 *pol* gene dataset encompassing 35 subtypes and 30,708 sequences, including 7,178 from Asia and 6,094 from Africa, demonstrating its robust generalizability.

However, this classification method has several limitations. Model performance depends on the quality of the annotated data. The model capability for discriminating rare subtypes with extremely limited sample sizes or URFs requires further validation with increased data accumulation. The model also lacks integration with relevant sociodemographic factors. Additionally, its ability to identify novel subtypes still requires its integration with more in-depth biological experiments.

In summary, we have generated a much-needed method for subtyping HIV-1 strains in China on the basis of specific gene fragments and full-genome sequences. This alignment-free and generalizable method can be directly applied to build a feature vector database of all circulating HIV-1 variants in China. Studies of its suitability for monitoring other rapidly evolving viruses are warranted.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the National Natural Science Foundation of China (Nos. 12031016 and 12531012), Interdisciplinary Construction of Bioinformatics and Statistics, and Academy for Multidisciplinary Studies, Capital Normal University.

doi: 10.46234/ccdcw2026.080

[#] Corresponding author: Hengjian Cui, hjcui@bnu.edu.cn.

[1] School of Science, Beijing University of Civil Engineering and Architecture, Beijing, China; [2] School of Intelligence Science and Technology, Beijing University of Civil Engineering and Architecture, Beijing, China; [3] Department of Mathematical Sciences, Tsinghua University, Beijing, China; [4] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China; [5] School of Mathematical Sciences, Capital Normal University, Beijing, China.
[&] Joint first authors.

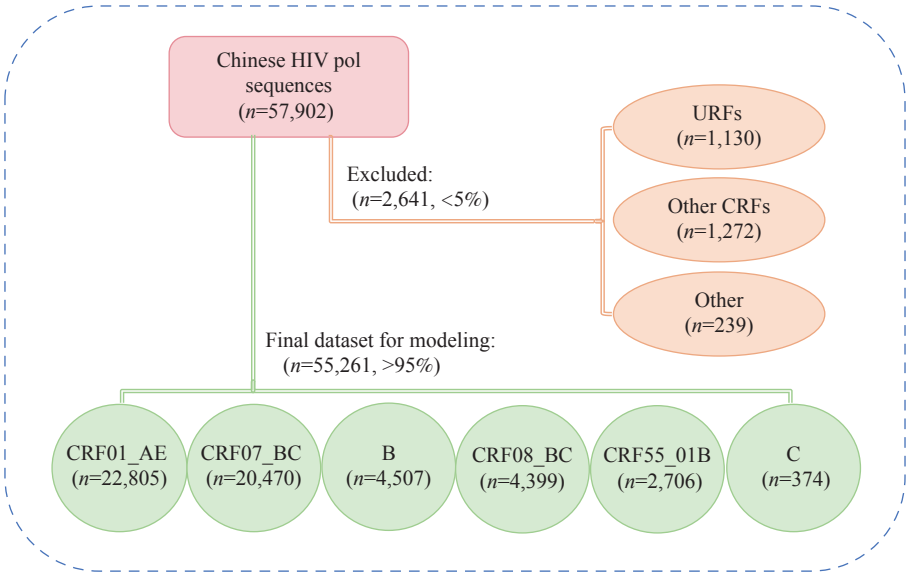
Submitted: January 11, 2026
Accepted: March 25, 2026
Issued: April 17, 2026

REFERENCES

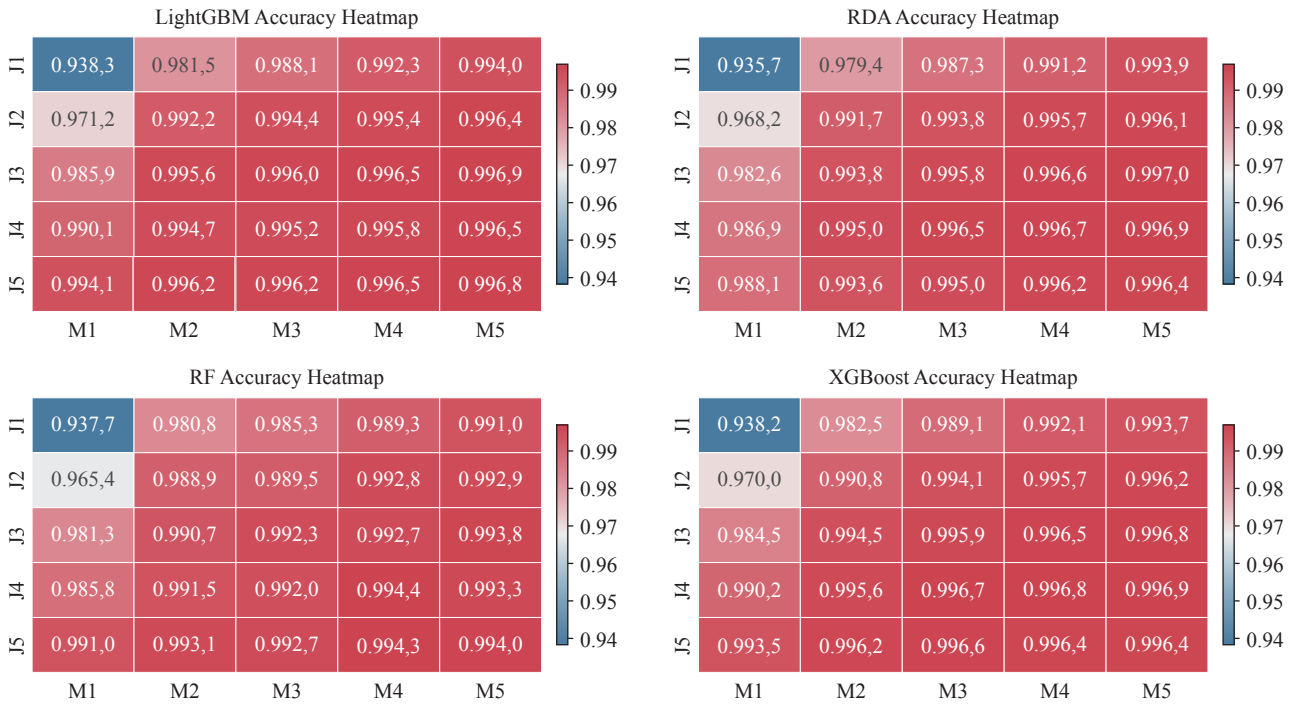
1. UNAIDS. 2025 Global AIDS Update. Geneva: Joint United Nations Programme on HIV/AIDS. 2025. https://www.unaids.org/en/resources/documents/2025/2025-global-aids-update. [2025-7-10]
2. Désiré N, Cerutti L, Le Hingrat Q, Perrier M, Emler S, Calvez V, et al. Characterization update of HIV-1 M subtypes diversity and proposal for subtypes A and D sub-subtypes reclassification. Retrovirology 2018;15(1):80. https://doi.org/10.1186/s12977-018-0461-y.
3. Liu X, Wang D, Hu J, Song C, Liao, LJ, Feng Y, et al. Changes in HIV-1 subtypes/sub-subtypes, and transmitted drug resistance among ART-naïve HIV-infected individuals—China, 2004-2022. China CDC Wkly 2023;5(30):664 – 71. https://doi.org/10.46234/ccdcw2023.129.
4. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, et al. Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis. Lancet Infect Dis 2019;19(2):143 – 55. https://doi.org/10.1016/S1473-3099(18)30647-9.
5. Hu L, Li ZF, Tang ZH, Zhao C, Zhou X, Hu PW. Effectively predicting HIV-1 protease cleavage sites by using an ensemble learning approach. BMC Bioinformatics 2022;23(1):447. https://doi.org/10.1186/s12859-022-04999-y.
6. He L, Sun SY, Zhang QY, Bao XN, Li PK. Alignment-free sequence comparison for virus genomes based on location correlation coefficient. Infect Genet Evol 2021;96:105106. https://doi.org/10.1016/j.meegid.2021.105106.
7. Wade KE, Chen LH, Deng CT, Zhou G, Hu PZ. Investigating alignment-free machine learning methods for HIV-1 subtype classification. Bioinform Adv 2024;4(1):vbae108. https://doi.org/10.1093/bioadv/vbae108.
8. Jiang S, Li HQ, Zhang LWY, Mu WP, Zhang Y, Chen TJ, et al. Generic Diagramming Platform (GDP): a comprehensive database of high-quality biomedical graphics. Nucleic Acids Res 2025;53(D1):D1670 – 6. https://doi.org/10.1093/nar/gkae973.
9. Struck D, Lawyer G, Ternes AM, Schmit JC, Bercoff DP. COMET:

adaptive context-based modeling for ultrafast HIV-1 subtype identification. Nucleic Acids Res 2014;42(18):e144. https://doi.org/10.1093/nar/gku739.
10. Pineda-Peña AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, et al. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools. Infect Genet Evol 2013;19:337 – 48. https://doi.org/10.1016/j.meegid.2013.04.032.
11. Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. Clin Infect Dis 2006;42(11):1608 – 18. https://doi.org/10.1086/503914.
12. He L, Dong R, He RL, Yau SST. A novel alignment-free method for HIV-1 subtype classification. Infect Genet Evol 2020;77:104080. https://doi.org/10.1016/j.meegid.2019.104080.

SUPPLEMENTARY MATERIAL



SUPPLEMENTARY FIGURE S1. Screening and composition of the Chinese HIV-1 *pol* dataset.
Abbreviation: HIV=human immunodeficiency virus; URFs=unique recombinant forms; CRF=circulating recombinant form.



SUPPLEMENTARY FIGURE S2. Heatmap of parameter selection for the four classifiers.
Abbreviation: LightGBM=Light Gradient Boosting Machine; RDA=regularized discriminant analysis; RF=random forest; XGBoost=eXtreme Gradient Boosting.

SUPPLEMENTARY TABLE S1. Classification performance of SCCFV-RDA on Chinese full-length genome sequences.

Dataset	Recall	Precision	F1-score	Sample size
CRF01_AE	1	1	1	299
CRF07_BC	1	1	1	298
CRF08_BC	1	1	1	52
CRF103_01B	1	1	1	10
CRF140_0107	1	1	1	19
CRF55_01B	1	1	1	27
CRF85_BC	1	1	1	51
B	1	1	1	65

Abbreviation: SCCFV-RDA=subsequence correlation coefficient feature vector-regularized discriminant analysis; CRF=Circulating recombinant form.

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. 16 Apr. 17, 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