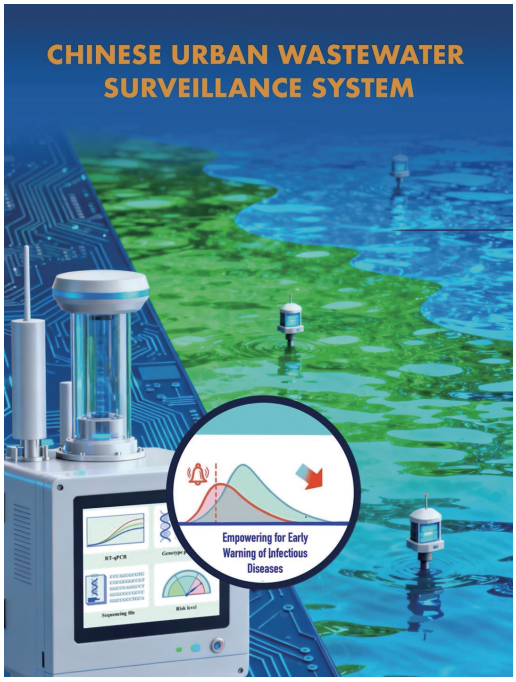
ISSN 2096-7071 (Print)
ISSN 2097-3101 (Online)
CN 10-1629/R1

CHINA CDC WEEKLY

Vol. 7 No. 51 Dec. 19, 2025

CCDC weekly

中国疾病预防控制中心周报



ISSN 2096-7071

China CDC Weekly

Editorial Board

Editorial Office

Cover Image: credited from the designer Jiaying Wang.

Recollections

Chinese Urban Wastewater Surveillance System for Early Warning of Infectious Diseases: Implementation and Efficacy — January 2023–June 2025

Jiayi Han[1]; Xia Li[1]; Shuxin Hao[1]; Xiao Zhang[1]; Liang Zhang[1]; Fuchang Deng[1]; Huihui Sun[1]; Yue Liu[1]; Yong Zhang[2]; Lin Wang[1]; Xiaoyuan Yao[1]; Lan Zhang[1,#]; Song Tang[3,#]; Xiaoming Shi[4,#]; Hongbing Shen[3]

ABSTRACT

Wastewater surveillance has emerged as a powerful tool for public health monitoring, particularly during disease outbreaks. This report documents China's pioneering establishment of the first nationwide comprehensive wastewater surveillance system with multi-scenario applications during the coronavirus disease 2019 (COVID-19) pandemic (January 2023–June 2025). The system integrated three components: national urban wastewater surveillance, inbound international aircraft wastewater surveillance, and pilot public health risk surveillance. This integrated framework demonstrated significant effectiveness in providing early warnings for COVID-19, polio, monkeypox, and other infectious disease outbreaks, while advancing variant tracking capabilities. These findings underscore the critical role of wastewater surveillance in augmenting existing public health infrastructure and improving outbreak detection capabilities. Implementing standardized protocols and developing collaborative networks will strengthen pandemic preparedness and enhance global public health resilience.

Infectious diseases continue to pose a formidable and evolving threat to global health security, as exemplified by the persistent emergence or resurgence of infectious disease outbreaks worldwide. Against this backdrop, establishing surveillance systems with broad-spectrum pathogen coverage and proactive early-warning capabilities has become essential for implementing targeted public health interventions. Environmental monitoring technologies have achieved large-scale applications over the past decade in tracking pathogen transmission dynamics and predicting outbreak risks by leveraging their unique strengths in unbiased population screening and early signal capture, driven by significant advancements in high-sensitivity molecular detection techniques. Among innovative surveillance strategies, wastewater surveillance based on wastewater-based epidemiology (WBE) theory has emerged as a particularly effective approach, characterized by its rapid, non-invasive, and scalable nature. This methodology provides objective, real-time, and cost-effective public health risk information within defined geographical areas (*1*). In response to the COVID-19 pandemic, 72 countries implemented wastewater surveillance for early outbreak detection and ongoing epidemiological monitoring (*2*). The Netherlands pioneered the first national wastewater surveillance framework for COVID-19 in early 2020, with sampling from 352 wastewater treatment plants (WWTPs) covering nearly all residents (*3*). Building on this advancement, the United States operationalized a nationwide wastewater surveillance system in September 2020, demonstrating dual functionality in both early outbreak detection and real-time data provision for localized public health decision-making. This system encompassed 1,567 sites at WWTPs or upstream locations (*4*). To complement existing population symptom-based surveillance (*5*), China launched the Chinese urban wastewater surveillance system (CWSS) to enhance national early monitoring and warning capabilities. Initiated in January 2023 under the auspices of the National Disease Control and Prevention Administration, the China CDC established this network across strategically selected cities nationwide to monitor and assess COVID-19 epidemic intensity, transmission dynamics, viral mutations, and other public health threats, both domestically and internationally. Throughout its operation, the CWSS has contributed significantly to the surveillance, prevention, and control of COVID-19, polio, monkeypox, and other epidemics. This report summarizes the implementation and efficacy of the CWSS from January 2023 to June 2025.

EXPERIENCE AND FINDINGS

System Architecture and Components

The CWSS surveillance system integrated three complementary operational components: 1) National urban wastewater surveillance: a network spanning 833 WWTPs across 169 cities, implementing semi-weekly 24-hour composite sampling at the facility inlets for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) detection and monthly poliovirus monitoring at designated sites. 2) Inbound international aircraft wastewater surveillance: weekly grab sampling from international flights arriving at 44 cities, with priority given to routes originating from high-risk countries and regions to enable early detection of SARS-CoV-2 variants and monkeypox lineages. Additionally, 12 cities expanded their surveillance capabilities to include targeted monitoring of highly pathogenic agents, specifically *Ebola virus*, *Marburg virus*, *Yellow fever virus*, MERS-CoV, and *Dengue virus*. 3) Urban wastewater public health risk surveillance: 13 pilot cities established 169 surveillance points that combined 24-hour composite sampling from WWTPs with sampling from strategically selected urban locations (e.g., hospitals, wet markets, and residential clusters) to simultaneously detect and quantify *influenza virus*, *norovirus*, toxigenic *Vibrio cholerae*, *Salmonella* spp., antimicrobial resistance genes (ARGs), pharmaceuticals related to infectious diseases (PRID, including antipyretics, analgesics, antibiotics, and psychotropics), and emerging threats such as Disease X. This comprehensive approach enabled robust population-level epidemiological trend analysis and antimicrobial resistance risk assessment. The CWSS network continues to expand, with the spatial distribution of surveillance sites illustrated in Figure 1.

Operational Workflow and Mechanisms

The CWSS received primary funding through central government appropriations, with local CDC departments coordinating sample collection, physicochemical analysis, and laboratory testing. Wastewater samples underwent enrichment and concentration procedures before reverse transcription polymerase chain reaction (RT-PCR) testing. Samples meeting eligibility criteria (Ct≤34) subsequently proceeded to genomic sequencing. Comprehensive laboratory outputs — including monitoring site information, physicochemical indicators, qualitative and quantitative viral nucleic acid results, and sequencing data — were uploaded to the China CDC surveillance platform for integrated analysis. Using SARS-CoV-2 as an exemplar, the surveillance system generated three key indicators: 1) the positive rate of viral nucleic acids in wastewater, reflecting the proportion of infected catchment areas; 2) the flow-weighted average of viral nucleic acid concentrations in wastewater, reflecting infection prevalence intensity;

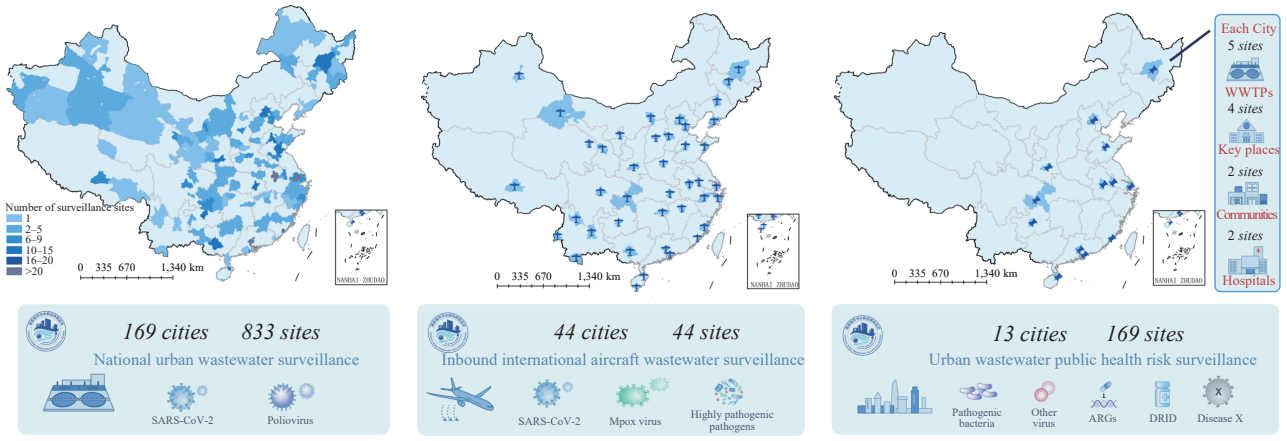


FIGURE 1. Geographic distribution of wastewater surveillance sites in the CWSS, January 2023–June 2025. (A) Geographic distribution of National urban wastewater surveillance sites; (B) Geographic distribution of Inbound international aircraft wastewater surveillance sites; (C) Geographic distribution of Urban wastewater public health risk surveillance sites.
Note: Each point represents an individual surveillance site. Key locations include bars, public bathhouses, retail supermarkets, and wet markets.
Map approval number: GS 京 (2025)2199 号.
Abbreviation: ARGs=antimicrobial resistance genes; PRID=pharmaceuticals related to infectious diseases; CWSS=Chinese Urban Wastewater Surveillance System.

and 3) viral genotype classification with composition ratios, identifying predominant pathogen strains. These indicators provided the foundation for epidemic trend analysis and informed targeted public health interventions. Upon detecting anomalous data fluctuations or emergent variants, the system initiated immediate verification protocols with local CDC departments for comprehensive risk assessment. Confirmed abnormal signals triggered formal risk evaluations, alert issuance, and recommendations for responsive countermeasures. The complete operational workflow of the surveillance system is illustrated in Figure 2.

System Performance and Impact

After two years of systematic implementation, the CWSS has matured into an integrated, agile, and highly responsive surveillance platform, generating over 200,000 data points that have proven instrumental in epidemic monitoring and early warning efforts. The continuous evolution of the CWSS has advanced methodological innovation in wastewater-based pathogen detection while simultaneously strengthening disease control capabilities across all tiers of the CDC organizational structure. Collectively, these advancements have established a robust foundation for rapid and effective responses to both current and emerging infectious disease threats.

Forecasting infectious disease epidemic trends. The CWSS has emerged as a critical tool for real-time monitoring of epidemiological dynamics and providing proactive early warnings for SARS-CoV-2, poliovirus, monkeypox virus, influenza, norovirus, and other pathogens. Using SARS-CoV-2 as a representative case, the CWSS employed a calibrated baseline methodology to establish risk thresholds based on quantitative wastewater pathogen load data, enabling scientific anomaly identification and comprehensive risk assessment. The spatiotemporal patterns of SARS-CoV-2 nucleic acid concentrations in urban wastewater revealed trajectories corresponding to four epidemic waves occurring in May and August 2023, and March and July 2024. These wastewater-derived patterns not only mirrored the peak incidence periods identified through conventional population-based surveillance but also provided the earliest objective signals of epidemic attenuation. Continuous, high-resolution spatiotemporal monitoring demonstrated that

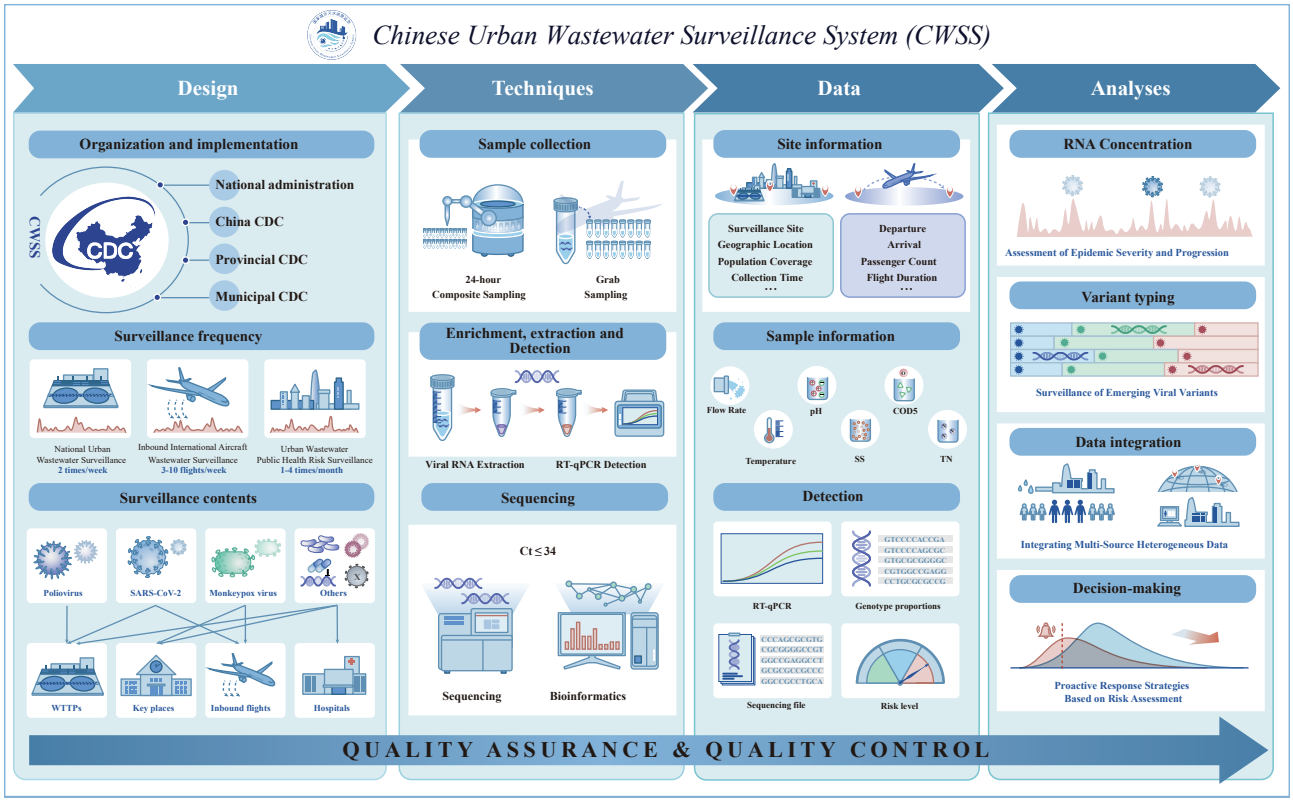


FIGURE 2. Chinese urban wastewater surveillance flowchart.
Abbreviation: SS=suspended solids; COD=chemical oxygen demand; TN=total nitrogen.

fluctuations in SARS-CoV-2 nucleic acid concentrations in wastewater exhibited statistically significant correlations with epidemiological trends observed in population surveillance findings (R^2=0.47). Correlation analyses revealed that wastewater surveillance could anticipate shifts in epidemic trends by 2–3 days relative to traditional population monitoring in Beijing (R^2=0.70–0.76) (*6*). The integration of these quantitative concentration measurements into surveillance frameworks has substantially enhanced the predictive and early warning capacities of public health authorities, enabling the implementation of precisely targeted and timely intervention strategies.

Timely warning of virus variants. Wastewater surveillance enabled comprehensive identification of viral variants, providing critical real-time insights into viral evolution, transmission dynamics, and potential outbreak trajectories. For example, in February 2023, municipal WWTPs successfully detected the *XBB.1.5* variant of SARS-CoV-2 — an internationally circulating strain not yet identified through domestic clinical surveillance — at a regional treatment facility. This environmental detection preceded the first local clinical case report by three days, immediately triggering targeted epidemiological investigations and intensified sentinel surveillance. The 72-hour lead time proved critical for implementing containment strategies and mitigating community transmission risk. Through strategic monitoring of high-risk venues (e.g., bars, clubs, and bathhouses) across selected pilot cities, wastewater surveillance detected monkeypox virus (subtype C1 Ⅱ b) on three occasions at a male bathing facility, successfully pinpointing a localized cluster and informing targeted public health interventions (*7*). Furthermore, proactive surveillance for novel pathogens *via* genome sequencing in wet market and community wastewater led to the identification of A/H5N6 avian influenza virus, which had not been detected in local human surveillance prior to this investigation. This pivotal finding immediately activated public health risk notification mechanisms, providing essential supplemental data to both human health and agricultural monitoring programs. Additionally, toxigenic *Vibrio cholerae* was isolated from hospital wastewater, including a strain of the O1 serogroup, prompting comprehensive population-based epidemiological investigations and case tracing efforts. Following expansion of the surveillance system, the number of poliovirus strains isolated in the past four months doubled compared to the cumulative total from previous years. This substantial increase underscored the enhanced sensitivity of wastewater-based surveillance over traditional case-based methods, thereby elevating both the predictive power and early warning capacity of infectious disease monitoring systems.

Promoting standardization of monitoring techniques. To address methodological challenges in SARS-CoV-2 wastewater surveillance, the CWSS project team systematically developed, validated, and implemented three enrichment and concentration methods: polyethylene glycol (PEG) precipitation, aluminum salt precipitation, and centrifugation with ultrafiltration. These protocols were published in March 2022 as the "Standard Method for Enrichment and Nucleic Acid Detection of SARS-CoV-2 in Sewage" (*8*), establishing a unified technological foundation for nationwide SARS-CoV-2 wastewater monitoring. Throughout implementation, the team continuously refined enrichment, concentration, nucleic acid detection, and variant identification protocols, achieving substantial improvements in detection sensitivity, viral recovery rates, and analytical accuracy (*9*). A comprehensive standardized operational framework was developed through iterative expert consultations and integration of field-derived empirical evidence. This structured framework encompassed nine core operational components: sample collection protocols, biohazard transport logistics, laboratory analytical procedures, data reporting architecture, quality assurance protocols, data analysis methods, risk stratification algorithms, bioinformatics pipelines, and emergency response cascades.

Developing an integrated data management system. A nationwide wastewater-based pathogen surveillance information system has been established to facilitate end-to-end digital management of data acquisition, transmission, and analysis through a dual-coding framework (site-specimen ID mapping). This platform enables systematic aggregation across three critical surveillance domains: 1) pathogenic profiles (including viral load quantification and variant subtyping), 2) pharmaceutical residues, and 3) antimicrobial resistance gene (ARG) profiling. By integrating wastewater pathogen monitoring data with clinical case reports and multi-source factors influencing disease outbreaks — including meteorological and environmental variables — the system employs epidemiological, statistical, and ecological analytical approaches to predict and provide early warnings of

potential infectious disease outbreaks through multi-dimensional datasets. The architecture incorporates an automated validation rule repository with a three-tier verification protocol (field-, provincial-, and national-level) to enable real-time dynamic tracking and accurate analysis of core metrics, such as spatiotemporal distribution of key variants and longitudinal trends in viral concentrations. This integrated system delivers rapid, sensitive, and intelligent monitoring capabilities that encompass data management, risk assessment, and outbreak prediction.

Advancing testing capacity in CDC organizations. The CWSS has substantially strengthened provincial and municipal CDC laboratory capabilities through a multifaceted capacity-building strategy encompassing specialized training programs, collaborative research initiatives, and nationwide proficiency testing with blinded wastewater samples. This systematic framework has enhanced pathogen detection sensitivity, upgraded analytical instrumentation infrastructure, and established wastewater surveillance as a cornerstone of infectious disease monitoring systems. The platform has further optimized cross-source data integration, refined early warning algorithms, and progressively enhanced predictive modeling capabilities for priority infectious diseases through continuous wastewater pathogen surveillance. To ensure sustained operational excellence and provide strategic guidance, the CWSS established the National Urban Wastewater Monitoring Project Expert Advisory Committee, comprising 13 distinguished scholars representing diverse disciplines including environmental health, microbiology, infectious disease epidemiology, bioinformatics, and biostatistics.

DISCUSSION

The effectiveness of infectious disease early warning systems fundamentally depends on their sensitivity and timeliness in detecting unusual increases or clusters of infections (*10*). In this regard, wastewater surveillance has emerged as an indispensable and transformative epidemiological tool. This approach excels in identifying viral variants and providing early warnings, thereby effectively monitoring population infection trends and epidemic inflection points. By generating high-frequency, real-time data, it supports evidence-based public health decision-making. Compared to traditional individual surveillance, wastewater surveillance reduces reliance on clinical case detection, circumvents the high costs of large-scale screening, and alleviates pressure on healthcare resources. Compared to wastewater surveillance systems in other countries, CWSS integrates municipal WWTPs, international flights, and high-risk venues into a cohesive multi-scenario surveillance framework. This integrated design addresses common fragmentation issues in multi-jurisdictional settings. The system's ability to track multiple priority pathogens simultaneously enhances surveillance efficiency. Additionally, the standardized technical protocols underpinning CWSS have been rigorously validated in over 120 cities, demonstrating a replicable and scalable model for broader implementation. Given its proven utility, the CWSS network continues to expand, comprising 1,076 monitoring sites as of June 2025, a scale that strengthens its capacity for monitoring emerging and re-emerging infectious diseases across diverse settings.

However, the findings in this report are subject to several key limitations. First, monitoring processes need further optimization in terms of standardization and normalization. Temporal variability in disease prevalence and inconsistencies in sampling, methodology, and equipment hinder cross-regional and temporal comparisons of pathogen concentrations in wastewater. Second, multi-pathogen detection requires technological innovation, with current methods focusing on known pathogens and lacking sensitivity. High-throughput detection is underdeveloped, and identifying unknown pathogens remains in its early stages. Third, data analysis and early warning capabilities need improvement. The interdisciplinary nature of wastewater surveillance creates challenges in obtaining comprehensive data, and advanced technologies like AI and machine learning are underutilized. Effective guidelines for integrating diverse data sources for early warning are needed.

Building on CWSS's current foundation, future strategic developments will focus on seven key priorities: 1) Standardizing sampling and testing protocols for consistent data across sites and periods; 2) Enhancing surveillance at critical points of entry (e.g., airports, hospitals) to detect and contain outbreaks; 3) Implementing region-specific protocols that prioritize local pathogens and public event impacts; 4) Advancing multi-pathogen detection for rapid, sensitive surveillance in wastewater; 5) Incorporating AI-driven, dynamic Bayesian networks and machine learning for predictive early warning and risk classification, and for optimizing thresholds;

6) Addressing technical bottlenecks (e.g., RNA loss) and integrating next-generation sequencing (NGS) technologies for real-time pathogen detection; 7) Balancing public health and privacy with a governance framework ensuring voluntary interventions rather than individual tracking. These improvements will strengthen China's predictive and early warning capabilities for infectious diseases, supporting a robust, responsive monitoring system.

CONCLUSIONS

Wastewater surveillance represents a critical and transformative tool for public health early warning. The CWSS framework provides a model that is efficient, flexible, fair, operational, sustainable, and policy-relevant, by virtue of its integrated multi-scenario design and capacity for simultaneous multi-pathogen tracking, making it suitable for nationwide implementation. Therefore, advancing and institutionalizing such systems is essential to substantially strengthen predictive capabilities and early warnings for major infectious diseases. This integrated approach offers fundamental scientific support for establishing a robust, intelligent, multi-point-triggered, and responsive national infectious disease monitoring and early warning system in China.

Conflicts of interest: No conflicts of interest.

Acknowledgments: The invaluable support and contributions of the local CDC organizations in each surveillance city, whose dedicated efforts were essential to the successful implementation of this nationwide surveillance system.

Funding: This research was supported by the National Urban Wastewater Priority Infectious Diseases Pathogen Surveillance Program, the National Key Research and Development Program (2023YFC3041300) from the Ministry of Science and Technology of the People's Republic of China, the Key Project of the Capital's Funds for Health Improvement and Research (2022-1G-4231), the Science and Technology Special Fund of Hainan Province (No. ZDYF2025SHFZ061), the Talent Support Program of Guangdong Provincial Center for Disease Control and Prevention (Z-QT2024D374), and the Young Scholar Science Foundation of China CDC (No. 2024A203).

doi: 10.46234/ccdcw2025.267

[#] Corresponding authors: Xiaoming Shi, shixm@chinacdc.cn; Song Tang, tangsong@nieh.chinacdc.cn; Lan Zhang, zhanglan@nieh.chinacdc.cn.

[1] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China; [2] National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China; [3] School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China; [4] Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine, Beijing, China.

Submitted: August 19, 2025
Accepted: December 02, 2025
Issued: December 19, 2025

REFERENCES

1. Kirby AE, Walters MS, Jennings WC, Fugitt R, LaCross N, Mattioli M, et al. Using wastewater surveillance data to support the COVID-19 response — United States, 2020–2021. MMWR Morb Mortal Wkly Rep 2021;70(36):1242 – 4. https://doi.org/10.15585/mmwr.mm7036a2.
2. COVIDPoops19. Summary of global SARS-CoV-2 wastewater monitoring efforts by UC Merced researchers. 2020. https://ucmerced.maps.arcgis.com/apps/opsdashboard/index.html#/c778145ea5bb4daeb58d31afee389082. [2024-12-12].
3. van Boven M, Hetebrij WA, Swart A, Nagelkerke E, van der Beek RF, Stouten S, et al. Patterns of SARS-CoV-2 circulation revealed by a nationwide sewage surveillance programme, the Netherlands, August 2020 to February 2022. Euro Surveill 2023;28(25):2200700. https://doi.org/10.2807/1560-7917.ES.2023.28.25.2200700.
4. Adams C, Bias M, Welsh RM, Webb J, Reese H, Delgado S, et al. The national wastewater surveillance system (NWSS): from inception to widespread coverage, 2020–2022, United States. Sci Total Environ 2024;924:171566. https://doi.org/10.1016/j.scitotenv.2024.171566.
5. CHINA CDC. The National Epidemic of SARS-CoV-2 Infections [EB/OL].(2023-01-15)[2025-11-06]. https://www.chinacdc.cn/jksj/xgbdyq/
6. Chen WX, An W, Wang C, Gao Q, Wang CZ, Zhang L, et al. Utilizing wastewater surveillance to model behavioural responses and prevent healthcare overload during "Disease X" outbreaks. Emerg Microbes Infect 2025;14(1):2437240. https://doi.org/10.1080/22221751.2024.2437240.
7. Xu J, Liu C, Zhang Q, Zhu HN, Cui F, Zhao ZQ, et al. The first detection of mpox virus DNA from wastewater in China. Sci Total Environ 2024;932:172742. https://doi.org/10.1016/j.scitotenv.2024.172742.
8. Zhang L, Huang X, Yang M. Methods for enrichment and nucleic acid detection of SARS-CoV-2 in sewage. Chin J Prev Med 2022;56(7):891 – 6. https://doi.org/10.3760/cma.j.cn112150-20220420-00396.
9. Du C, Peng YJ, Lyu ZQ, Yue ZJ, Fu YL, Yao XJ, et al. Early detection of the emerging SARS-CoV-2 BA.2.86 lineage through wastewater surveillance using a mediator probe PCR assay — Shenzhen City, Guangdong Province, China, 2023. China CDC Wkly 2024;6(15):332-8. http://dx.doi.org/10.46234/ccdcw2024.063.
10. Yang WZ, Lan YJ, Lyu Y, Leng ZW, Feng LZ, Lai SJ, et al. Establishment of multi-point trigger and multi-channel surveillance mechanism for intelligent early warning of infectious diseases in China. Chin J Epidemiol 2020;41(11):1753 – 7. https://doi.org/10.3760/cma.j.cn112338-20200722-00972.

Vital Surveillances

Environmental Health Literacy Prevalence and Profiles among Shanghai Residents — Shanghai, China, 2020–2024

Fengchan Han[1]; Ling Tong[1]; Duo Wang[1]; Zheng Wu[1]; Hailei Qian[1]; Yewen Shi[1]; Chunyang Dong[1]; Feier Chen[1]; Chen Wu[1]; Yangyang Ren[1]; Mingjing Xu[1]; Mengshuang Liu[1]; Aimin Du[1]; Zhenni Zhu[1]; Yi He[1]; Shaofeng Sui[1]; Tian Chen[2,#]; Jianghua Zhang[1,#]

ABSTRACT

Introduction: In 2019, the Chinese State Council launched the "Healthy China Initiative (2019–2030)", establishing explicit targets for residents' environmental and health literacy (EHL): reaching to 15% by 2022, to 25%, and over 2030. To identify knowledge gaps and guide targeted interventions, Shanghai implemented five consecutive EHL surveys between 2020 and 2024.

Methods: We employed a multi-stage random sampling design across five cross-sectional surveys. Associations with EHL levels were examined using χ^2 tests, one-way analysis of variance, generalized linear models, and multivariate logistic regression analyses.

Results: Among 11,220 residents aged 15–69 years assessed using the Core Questionnaire for Assessing the EHL of Chinese Residents, mean EHL scores demonstrated steady improvement. Scores increased from 55.28±15.64 points in 2020 to 61.77±15.92 points (2021), 62.13±17.14 points (2022), 62.03±16.97 points (2023), and 63.14±18.21 points (2024) (P<0.001). The proportion achieving adequate EHL (≥70 points) increased correspondingly, with age-adjusted rates rising from 18.78% in 2020 to 30.18% (2021), 33.22% (2022), 33.84% (2023), and 42.88% (2024). Among the three primary dimensions, knowledge showed the greatest improvement, increasing from 7.12% to 39.93%. Participants surveyed in 2024 had 3.50-fold higher odds of achieving adequate EHL compared with those in 2020 (odds ratio=3.50; 95% confidence interval: 3.07, 4.00).

Conclusions: Although educational attainment remained the primary determinant of EHL, targeted public health education campaigns significantly improved EHL among Shanghai residents between 2020 and 2024.

Rapid population growth and expanding human activities are intensifying environmental degradation, depleting natural resources, and increasingly threatening human health (*1*). The World Health Organization estimates that 24% of all global deaths and 28% of deaths among children under 5 years of age are attributable to modifiable environmental hazards, most of which could be prevented through the establishment of healthier environments (*2*).

Consequently, environmental health literacy (EHL) has evolved beyond mere recognition of exposure-disease linkages to encompass a comprehensive understanding that includes valuing intact ecosystems for human well-being, mastering knowledge related to ecological protection and health risk prevention, and adopting sustainable, healthy lifestyles (*3*–*5*). These competencies contribute to environmental protection and preservation while promoting individual health. As the burden of environment-related diseases escalates, EHL is attracting increasing global attention (*1*–*5*).

In July 2020, the Ministry of Ecology and Environment replaced the 2013 "Citizen EHL (Trial)" with the updated "Chinese Citizens' Ecological EHL," establishing a national framework for disseminating environmental health knowledge, attitudes, and skills (*6*–*7*). Building on this foundation, the Chinese State Council initiated the "Healthy China Initiative (2019–2030)," which established explicit targets for residents' EHL: an increase to 15% by 2022 and 25% by 2030. To identify gaps in environmental health knowledge dissemination and facilitate targeted improvements, Shanghai municipal institutions related to environmental health launched the "Environmental Health Literacy Survey and Improvement Program." A baseline EHL survey stratified by gender and age was conducted, followed by theory-driven interventions implemented district-wide. Multi-channel

dissemination strategies — including print materials, social media, subway carriage displays, and gamified quizzes — delivered environmental health education. Annual follow-up surveys tracked EHL progress, with the goal of surpassing national 2022 and 2030 benchmarks.

METHODS

Between 2020 and 2024, we conducted five annual cross-sectional surveys in Shanghai. Eligible participants included residents aged 15–69 years who had resided in the study area for at least six months during the previous year; we excluded individuals living in group quarters, such as student dormitories or employee housing (*6*–*7*). To ensure representativeness and comparability across survey years, we employed multi-stage cluster random sampling (*6*–*7*).

Initially, we calculated the minimum sample size for each stratum using equation (1):

$$n_{min} = [z_{\alpha}^{2} p(1-p)/(p \times re)^{2}] \times deff \quad (1)$$

where z=1.96 (α=0.05), p=0.5 [environmental health literacy (EHL) prevalence assumed in the absence of prior data], re=0.15 (relative error), and $deff$=1.5 (design effect).

Next, we inflated this minimum to the final target sample size as equation (2):

$$N = n_{\min} \times (\text{product of stratification factors}) \times (1 + \text{refusal rate}) \quad (2)$$

Stratification was based on gender (male, female) and survey areas (urban, suburban), resulting in four strata (2×2). Assuming a 14% refusal rate, the required sample size was at least 1,127 participants each year. The 2020 Shanghai survey followed the 2013 and 2017 trial guidelines and oversampled older adults to reflect the city's aging demographic structure. However, to align with the 2022 evaluation deadline of the Healthy China Initiative, the Ministry of Ecology and Environment issued updated national documents in 2021 (trial version) and 2022 (final version of the "Survey Protocol for Residents' Environmental Health Literacy"). Beginning in 2021, Shanghai adopted the national protocol, which employed a younger age distribution standard. This methodological shift resulted in demographic differences between the 2020 baseline and subsequent survey years (2021–2024), though the 2021–2024 surveys maintained internal demographic consistency.

The questionnaire assessed socio-demographic characteristics and incorporated the 47-item Core Questionnaire for Assessing the EHL of Chinese Residents, developed by the Ministry of Ecology and Environment. Sampling procedures and scoring methods followed the supplemental materials and protocols previously reported (*6*–*10*). Across the five survey rounds, we obtained 11,672 questionnaires and retained 11,220 valid responses, yielding a response rate of 96.13%.

Statistical Analysis

All statistical analyses were conducted using R version 4.2.2 (R Foundation for Statistical Computing, Auckland, New Zealand). Environmental health literacy scores are reported as mean±standard deviation with interquartile ranges (P25–P75). We assessed between-group differences in continuous scores using independent t-tests or one-way analysis of variance, while categorical differences in EHL levels were evaluated with χ^2 tests. To identify factors associated with EHL, we employed generalized linear models and multivariate logistic regression analyses.

RESULTS

Demographic Characteristics

We analyzed 11,220 valid questionnaires distributed across the five survey years: 3,720 in 2020, 3,180 in 2021, and 1,440 in each of 2022, 2023, and 2024. Table 1 presents the demographic characteristics of respondents, including areas, gender, age, ethnicity, education level, occupation, and per-capita monthly income. Significant differences across survey years were observed for age, education, occupation, and income distributions (P<0.05).

EHL Scores

The overall mean EHL score was 59.87±16.76, falling below the 70-point threshold for adequate literacy. No significant differences emerged by survey areas (urban *vs.* suburban), gender, or ethnicity (P>0.05). Compared with the 2020 baseline (55.28±15.64), mean scores demonstrated consistent improvement: 61.77±15.92 in 2021, 62.13±17.14 in 2022, 62.03±16.97 in 2023, and 63.14±18.21 in 2024 (P<0.001). EHL scores followed an inverted-U pattern across age groups, peaking in middle age before declining (P<0.001). In contrast, scores increased monotonically with both education and income levels (both P<0.001) and varied significantly by occupation

TABLE 1. Demographic characteristics of the study population.

Variables	Factors	2020		2021		2022		2023		2024		χ^2	*P*
		N (3,720)	(%)	*N* (3,180)	(%)	*N* (1,440)	(%)	*N* (1,440)	(%)	*N* (1,440)	(%)		
Areas	Suburban	1,973	53.0	1,620	50.9	720	50.0	720	50.0	720	50.0	7.573	0.109
	Urban	1,747	47.0	1,560	49.1	720	50.0	720	50.0	720	50.0		
Gender	Male	1,873	50.4	1,548	48.7	720	50.0	720	50.0	720	50.0	2.097	0.718
	Female	1,847	49.6	1,632	51.3	720	50.0	720	50.0	720	50.0		
Age, years	15–17	30	0.8	315	9.9	96	6.7	96	6.7	96	6.7	497.922	0.000
	18–34	915	24.6	964	30.3	432	30.0	432	30.0	432	30.0		
	35–49	1,023	27.5	1,017	32.0	432	30.0	432	30.0	432	30.0		
	50–69	1,752	47.1	884	27.8	480	33.3	480	33.3	480	33.3		
Ethnicity	Han nationality	3,687	99.1	3,156	99.3	1,432	99.4	1,430	99.3	1,436	99.7	6.088	0.193
	Others	33	0.9	24	0.7	8	0.6	10	0.7	4	0.3		
Education	Primary school and below	357	9.6	129	4.1	89	6.2	7	0.5	9	0.6	1,143.22	<0.000
level	Junior high school	917	24.7	718	22.6	310	21.5	75	5.2	54	3.8		
	Senior high school	895	24.1	874	27.5	662	46.0	632	43.9	594	41.3		
	Vocational college/undergraduate	1,493	40.1	1,413	44.4	346	24.0	681	47.3	749	52.0		
	Postgraduate and above	58	1.5	46	1.4	33	2.3	45	3.1	34	2.4		
Occupation	Farmer	263	7.1	122	3.8	94	6.5	69	4.8	67	4.7	3,839.381	0.000
	Labor in city	1,850	49.7	1,647	51.8	693	48.1	608	42.2	616	42.8		
	Civil servants and leaders	242	6.5	117	3.7	70	4.9	56	3.9	48	3.3		
	Student	91	2.5	437	13.7	152	10.6	177	12.3	182	12.6		
	Retiree	1,071	28.8	593	18.7	284	19.7	318	22.1	307	21.3		
	Others	203	5.4	264	8.3	147	10.2	212	14.7	220	15.3		
Per capita	<5,500	1,098	29.5	1,058	33.27	/	/	470	32.6	395	27.4	2,652.866	0.000
Monthly	5,500–12,999	1,927	51.8	1,415	44.5	/	/	767	53.3	717	49.8		
Income	13,000–20,999	375	10.1	467	14.69	/	/	184	12.8	223	15.5		
(CNY)*	>21,000	320	8.6	240	7.55	/	/	19	1.3	105	7.3		

Abbreviation: CNY=Chinese Yuan.
* This variable was not investigated in 2022.

(*P*<0.001). Post-hoc analyses revealed no significant differences between laborers and other occupational groups, and between civil servants and students (*P*>0.05). Detailed results are presented in Supplementary Figure S1 (available at https://weekly.chinacdc.cn).

EHL Levels

Adequate EHL was defined as a total score of ≥70 points. Across the five survey waves (2020–2024), 30.74% of participants achieved this threshold, with the proportion increasing steadily: 18.78% in 2020, 33.24% in 2021, 36.18% in 2022, 36.67% in 2023, and 44.72% in 2024 (*P*<0.001). However, Table 1 reveals a substantial shift in age distribution between 2020 and subsequent years. The 2020 survey reflected Shanghai's actual population structure (aged 15–69 years), whereas the 2021–2024 surveys adopted a younger national standard. To preserve validity and isolate temporal trends from demographic shifts, we age-standardized the 2021–2024 estimates to Shanghai's 2020 age composition. Age-adjusted EHL demonstrated significant improvement: 18.78% (2020), 30.18% (2021), 33.22% (2022), 33.84% (2023), and 42.88% (2024). The lower age-standardized rates compared with crude rates reflect the down-weighting of younger respondents when applying the 2020 reference population, which contained a higher proportion of elderly individuals. Urban residents demonstrated higher literacy than their suburban counterparts (32.00% *vs.* 29.41%), while men achieved marginally higher rates than women (31.85% *vs.* 29.63%). Age exhibited an

inverted-U pattern, with literacy peaking at 40.41% among individuals aged 18–34 years before declining to 20.45% in the 50–69 years age group. Education displayed a clear dose–response gradient: only 13.18% of participants with primary schooling or less achieved adequate EHL, compared with 47.00% of those with postgraduate education (P<0.001). Complete demographic breakdowns are presented in Table 2.

Classification EHL level represents the proportion of participants whose score in a given domain achieved or exceeded 70% of the maximum possible score for that domain. The survey evaluated three first-level domains — basic concepts, basic knowledge, and basic skills — alongside six second-level domains: basic cognition, basic attitudes, fundamental concepts, scientific knowledge, basic behavior, and basic skills. Figure 1 presents the temporal trends for both first- and second-level domains across 2020–2024. Among the first-level domains (Figure 1A), basic concepts exhibited an initial increase followed by a temporary decline before rising again, whereas basic skills maintained consistently high levels with minimal

TABLE 2. Proportion of participants achieving adequate EHL by sociodemographic characteristics.

Variables	Factors	Sample size	Total EHL level (%)	χ^2	P
Total	Sum	11,220	30.74	/	/
Year	2020	3,720	18.78	435.443	0.000
	2021	3,180	33.24		
	2022	1,440	36.18		
	2023	1,440	36.67		
	2024	1,440	44.72		
Areas	Suburban	5,467	29.41	8.813	0.003
	Urban	5,753	32.00		
Gender	Male	5,581	31.85	6.519	0.011
	Female	5,639	29.63		
Age, years	15–17	633	39.91	371.542	0.000
	18–34	3,175	40.41		
	35–49	3,336	32.34		
	50–69	4,076	20.45		
Ethnicity	Han nationality	11,143	30.69	1.333	0.248
	Others	77	36.71		
Education	Primary school and below	591	13.18	543.914	0.000
Level	Junior high school	2,074	15.56		
	Senior high school	3,656	28.98		
	Vocational college/undergraduate	4,683	40.30		
	Postgraduate and above	216	47.00		
Occupation	Farmer	615	17.69	466.704	0.000
	Labor in city	5,414	29.40		
	Civil servants and leaders	533	45.65		
	Student	1,039	43.27		
	Retiree	2,573	20.39		
	Others	1,046	33.33		
Per capita	<5,500	2,157	18.48	173.473	0.000
Monthly	5,500–12,999	4,827	33.13		
Income	13,000–20,999	1,248	33.68		
(CNY)*	≥21,000	1,548	32.90		

Abbreviation: EHL=Environmental Health Literacy, CNY=Chinese Yuan.
* This variable was not investigated in 2022.

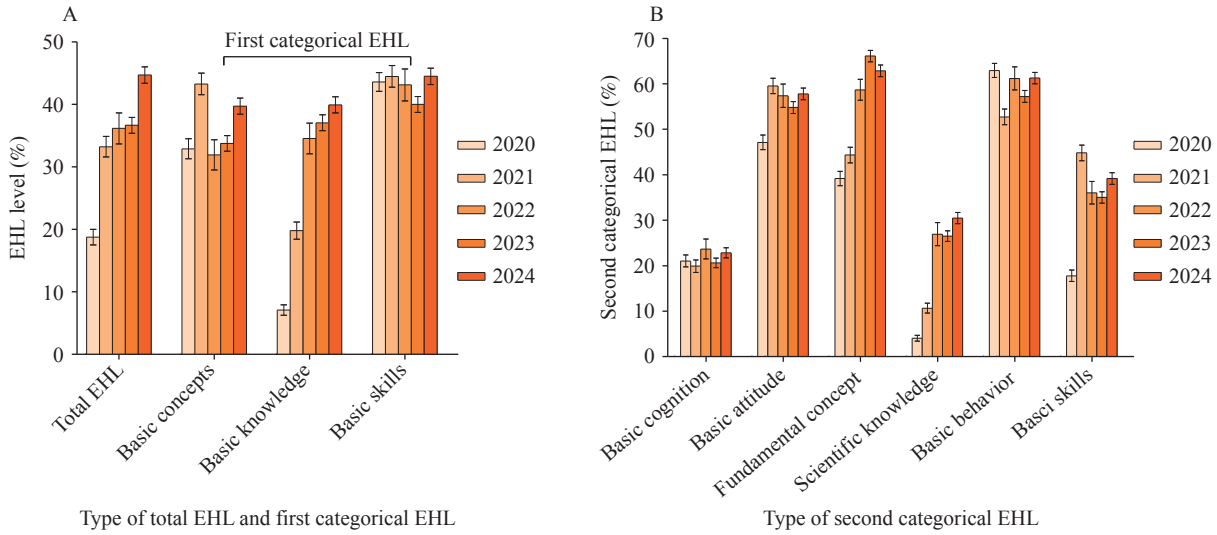


FIGURE 1. Proportion of Shanghai residents achieving adequate EHL, 2020–2024. (A) Overall EHL and its three first-level components; (B) Six second-level components.

Note: Error bars represent the standard error, calculated as $\sqrt{p(1-p)/n}$, in which p=EHL level and n=sample size.

Abbreviation: EHL=Environmental Health Literacy.

variation. In contrast, basic knowledge demonstrated the most substantial improvement, increasing markedly from 7.12% in 2020 to 19.81%, 34.58%, 37.08%, and 39.93% in 2021–2024, respectively. At the second-level (Figure 1B), basic cognition remained relatively stable near 20% throughout the study period, while scientific knowledge, fundamental concepts, and basic skills all showed considerable improvement compared with baseline measurements in 2020.

Determinants of EHL

EHL is influenced by survey year, survey area, gender, age, education, occupation and income. After adjusting for potential confounders, multivariate generalized-linear and logistic regression models were employed.

Compared with 2020, mean EHL in 2024 was 7.87 points higher (95% *CI*: 6.87, 8.87). Urban residents scored 1.78 points higher than suburban residents (95% *CI*: 1.07, 2.50). The youngest age group (15–34 years) outperformed the oldest (50–69 years) by 9.42 points (95% *CI*: 8.06, 10.79), and the highest-education group exceeded the lowest by 15.54 points (95% *CI*: 13.03, 18.05). Civil servants and senior managers scored 10.66 points higher than farmers (95% *CI*: 9.15, 12.18), while the middle-income band (CNY 13,000–21,000) scored 6.56 points above the lowest-income group (95% *CI*: 5.41, 7.71) (Figure 2A).

In the logistic model, participants in 2024 were 3.50 times more likely to achieve adequate EHL than those in 2020 (*OR*=3.50; 95% *CI*: 3.07, 4.00). Suburban residents had lower odds than urban residents (*OR*=0.89; 95% *CI*: 0.82, 0.96), as did women compared with men (*OR*=0.90; 95% *CI*: 0.83, 0.98). Relative to the youngest group, the oldest had markedly lower odds (*OR*=0.39; 95% *CI*: 0.33, 0.46), whereas the highest-education group demonstrated nearly six-fold greater odds (*OR*=5.85; *95% CI*: 4.08, 8.36) (Figure 2B).

DISCUSSION

By 2020, 18.78% of Shanghai residents had already surpassed the 2022 national target of 15% established by the Healthy China Initiative (2019–2030) (*11*). This baseline figure exceeded contemporaneous statistics from Shaanxi (17.6%) (*8*), Hubei (17.4%) (*10*), and the initial 2020 national survey (12.5%) (*12*), while closely approximating the second national survey conducted in 2022 (18.8%) (*13*). When compared with the 2022 national survey results, Shanghai residents demonstrated substantially higher overall EHL (36.18% *vs.* 18.8%) and superior performance in both basic knowledge (34.58% *vs.* 14.1%) and basic skills (43.14% *vs.* 25.7%) (*13*). However, performance in basic concepts lagged slightly behind national levels (31.94% *vs.* 36.1%). Among second-level domains, basic cognition (23.68%) and scientific knowledge (26.94%) remained the principal

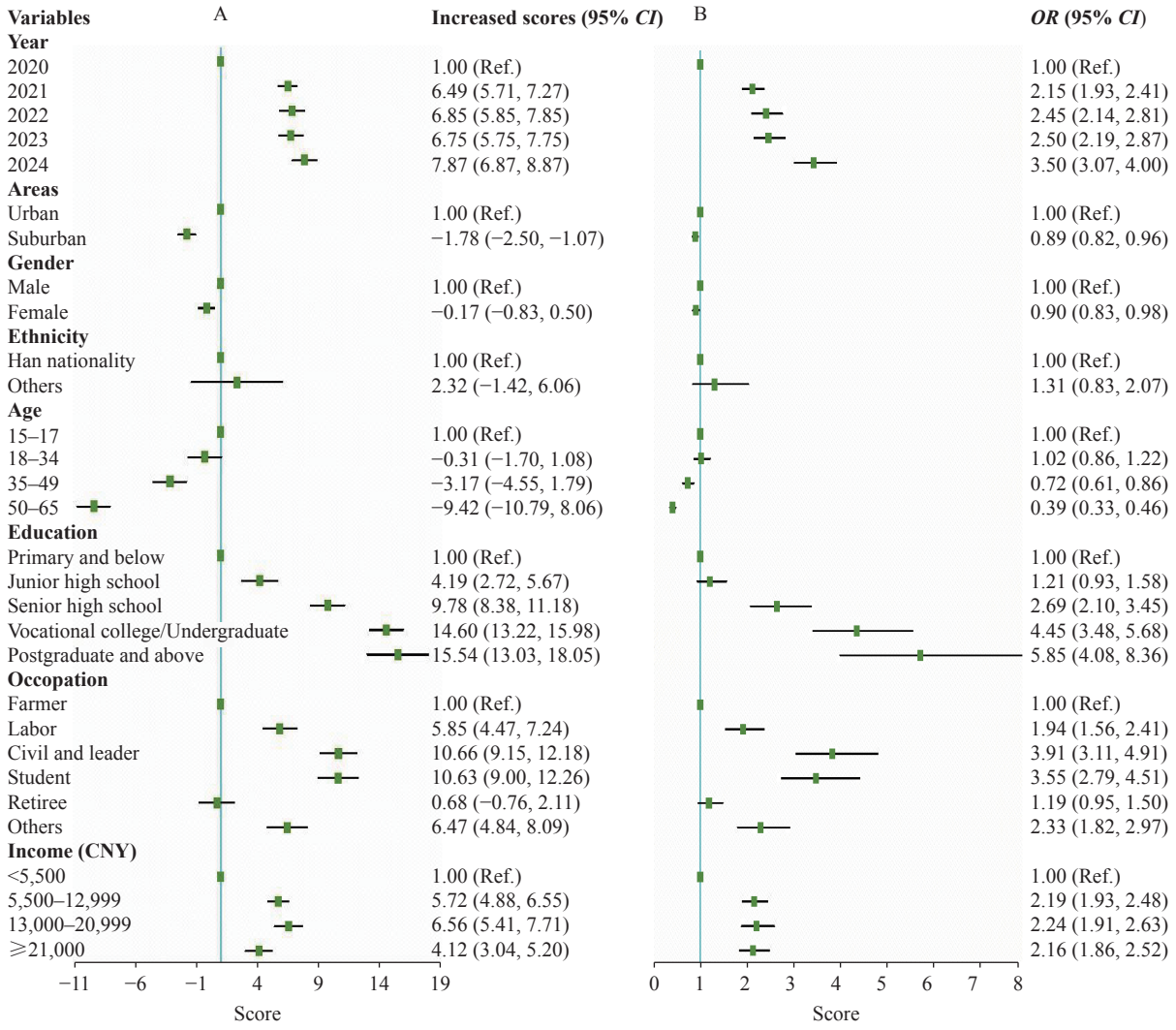


FIGURE 2. Determinants of EHL in Shanghai, 2020–2024. (A) Change in mean EHL score (generalized-linear model); (B) Odds of achieving adequate EHL (multivariate logistic regression).
Abbreviation: *CI*=confidence interval; EHL=Environmental Health Literacy; CNY=Chinese Yuan.

areas requiring improvement (*13*), though both approached the 2030 national threshold of 25% (*11*). Across occupational categories, farmers exhibited the lowest EHL levels, whereas civil servants and organizational leaders recorded the highest, reflecting underlying educational disparities. Although men achieved slightly higher scores than women (31.85% *vs.* 29.63%), this gender gap was substantially narrower than that observed in Shaanxi (25.0% *vs.* 11.5%) (*8*) or Hubei (20.6% *vs.* 15.8%) (*10*), likely attributable to higher educational attainment among women in Shanghai. Similarly, the suburban–urban disparity (29.4% *vs.* 32.0%) was less pronounced than that documented in Shaanxi (*8*), Hubei (*10*), or the 2022 national survey (*13*), indicating more balanced regional development in Shanghai.

Although demographic characteristics varied across survey years (Table 1), Spearman correlation analyses revealed moderate negative associations between age and education (r=–0.38) and weak positive associations between age and both occupation and income (r=0.14 and 0.17, respectively). These correlations demonstrate that age is interrelated with education, occupation, and income. Age standardization was therefore employed to control for these socioeconomic differences across survey years. Furthermore, these correlations indicate that older residents tend to have lower educational attainment, are more likely to work in agriculture, and earn less — explaining why education emerges as the dominant determinant of EHL. This finding mirrors earlier research demonstrating that low educational attainment predicts higher mortality, with income

serving as a key mediator (*14–15*); improving EHL may therefore help offset these risks by promoting healthier environments (*2*).

To sustain and further enhance EHL, Shanghai should prioritize health promotion and surveillance among vulnerable populations — specifically older adults, individuals with limited education or income, and agricultural workers — while focusing educational content on fundamental environment-health relationships and scientific knowledge regarding air quality, water safety, soil contamination, ocean health, biodiversity conservation, and climate change. Future EHL improvement initiatives should employ tailored dissemination strategies for different target audiences. For instance, short, accessible videos on environmental health topics can be delivered through social media platforms such as TikTok to reach older adults who regularly use these applications, whereas informational displays in subway stations and transit systems can effectively target commuters and office workers.

This study's principal strength lies in its five-year series of repeated cross-sectional surveys conducted in a large, representative sample, yielding actionable evidence for public health interventions. The primary limitation stems from the shift in age composition between the 2020 baseline and subsequent 2021–2024 surveys. Given Shanghai's older demographic structure, nationally age-standardized crude EHL rates for 2022–2024 overestimate city-level literacy. Consequently, locally age-standardized rates provide a more accurate measure of Shanghai-specific improvements in EHL. Additionally, unregistered renters were excluded from sampling; a targeted survey of this population is needed to fully assess the equity and reach of current health-promotion initiatives.

CONCLUSIONS

Using baseline data from 2020, we designed targeted interventions that prioritized suburban residents, older adults, and individuals with low income or education levels, delivering tailored content to each group. Core educational messages encompassed fundamental environment-health interactions, regulatory standards, water quality, environmental toxicants, air pollution, and climate change. We employed multimodal, audience-specific dissemination channels — including print media, newspapers, and paper-based materials for older adults — to address identified knowledge gaps. Five consecutive city-wide surveys demonstrated a significant monotonic increase in age-standardized EHL: 18.78% (2020), 30.18% (2021), 33.22% (2022), 33.84% (2023), and 42.88% (2024). The most substantial gains occurred in the first post-intervention year, with continued improvements observed in subsequent years. These findings substantiate the effectiveness of Shanghai's environmental health education campaigns.

Conflicts of interest: No conflicts of interest.

Acknowledgements: The field teams from the district-level Centers for Disease Control and Prevention and the community health service centers for their meticulous data collection. We are also grateful to all Shanghai residents who participated in these surveys.

Ethical statement: Approved by the Shanghai Municipal Center for Disease Control and Prevention (approval No. 2020-94).

Funding: Supported by the Key Projects in the Three-Year Plan of Shanghai Municipal Public Health System (2023–2025) (No. GWV1-11.1-39), the National Disease Control and Prevention Administration Talent Project of Field Epidemiological Investigation (No. Y2023-28B), and the Shanghai Municipal Health Commission Science and Youth Research Fund (No. 20234Y0297).

doi: 10.46234/ccdcw2025.269

[#] Corresponding authors: Tian Chen, chentian@scdc.sh.cn; Jianghua Zhang, zhangjianghua@scdc.sh.cn.

[1] Division of Health Risk Factors Monitoring and Control/State Environmental Protection Key Laboratory of Environmental Health Impact Assessment of Emerging Contaminants, Shanghai Municipal Center for Disease Control and Prevention (Shanghai Academy of Preventive Medicine), Shanghai, China; [2] Shanghai Preventive Medicine Association, Shanghai, China.

Submitted: September 26, 2025
Accepted: December 10, 2025
Issued: December 19, 2025

REFERENCES

1. Chowdhury R, Ramond A, O'Keeffe LM, Shahzad S, Kunutsor SK, Muka T, et al. Environmental toxic metal contaminants and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 2018;362:k3310. https://doi.org/10.1136/bmj.k3310.
2. World Health Organization. Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks. Geneva: World Health Organization; 2016. https://www.unhabitat-urbanhealth.org/download/preventing-disease-through-healthy-environments-a-global-assessment-of-the-burden-of-disease-from-environmental-risks/.
3. Burke TA, Cascio WE, Costa DL, Deener K, Fontaine TD, Fulk FA,

et al. Rethinking environmental protection: meeting the challenges of a changing world. Environ Health Perspect 2017;125(3):A43 – 9. https://doi.org/10.1289/EHP1465.
4. Maindal HT, Aagaard-Hansen J. Health literacy meets the life-course perspective: towards a conceptual framework. Glob Health Action 2020;13(1):1775063. https://doi.org/10.1080/16549716.2020.1775063.
5. Finn S, O'Fallon L. The emergence of environmental health literacy—from its roots to its future potential. Environ Health Perspect 2017;125(4):495 – 501. https://doi.org/10.1289/ehp.1409337.
6. Ministry of Ecological Environment of the People's Republic of China. Announcement on the release of the ecological environment and health literacy for Chinese citizens. 2020. https://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/202007/t20200727_791324.html. [2025-8-25]. (In Chinese).
7. Ministry of Ecology and Environment of the People's Republic of China. Technical guidelines for measuring citizens' environmental and health literacy (for trial implementation). 2017. https://www.mee.gov.cn/gkml/hbb/bgg/201706/W020170608498381052378.pdf. [2025-08-25]. (In Chinese).
8. Zhao Y, Sheng Y, Zhou JT, Wang H, Chilufya MM, Liu X, et al. Influencing factors of residents' environmental health literacy in Shaanxi province, China: a cross-sectional study. BMC Public Health 2022;22(1):114. https://doi.org/10.1186/s12889-022-12561-x.
9. Deng FD, Wen XY, Dong GQ, Wang XN, He HF, Zhu RX, et al. The environmental health literacy level was effectively improved of residents in Shaanxi Province, China, 2022. Front Public Health 2025;12:1499349. https://doi.org/10.3389/fpubh.2024.1499349.
10. Wang Q, Liu Q, Hou AL, Yu Y, Hu CH, Liang SW. Analysis on the current situation and influencing factors of environmental and health literacy of residents in Hubei province. Environ Sci Technol 2020;43(5):230 – 6. https://doi.org/10.19672/j.cnki.1003-6504.2020.05.031.
11. The State Council of the People's Republic of China. Action of Healthy China (2017–2030). 2019. https://baike.baidu.com/item/%E5%81%A5%E5%BA%B7%E4%B8%AD%E5%9B%BD%E8%A1%8C%E5%8A%A8%EF%BC%882019%E2%80%942030%E5%B9%B4%EF%BC%89/23543779. [2023-6-21]. (In Chinese).
12. Ministry of Ecology and Environment of the People's Republic of China. Results of the first survey of Chinese residents' environmental health literacy. 2020. http://www.mee.gov.cn/ywgz/fgbz/hjyjk/gzdt/202008/t20200810_793281.shtml. [2020-8-10]. (In Chinese).
13. Ministry of Ecology and Environment of the People's Republic of China. Results of the survey of Chinese residents' environmental health literacy in 2022. 2023. https://www.mee.gov.cn/ywgz/fgbz/hjyjk/gzdt/202311/t20231120_1056812.shtml. [2023-11-20]. (In Chinese).
14. Wang JF, Zhao ZP, Yang J, Ng M, Zhou MG. The association between education and premature mortality in the Chinese population: a 10-year cohort study. Lancet Reg Health West Pac 2024;47:101085. https://doi.org/10.1016/j.lanwpc.2024.101085.
15. IHME-CHAIN Collaborators. Effects of education on adult mortality: a global systematic review and meta-analysis. Lancet Public Health 2024;9(3):e155 – 65. https://doi.org/10.1016/S2468-2667(23)00306-7.

SUPPLEMENTAL MATERIALS

Sampling Method

Between 2020 and 2024, we conducted five annual cross-sectional surveys in Shanghai. Eligible participants were residents aged 15–69 years who had lived in the study area for at least six months during the previous year; we excluded individuals residing in group quarters, such as student dormitories or employee housing (*1–2*). To ensure representativeness and enable valid comparisons across survey years, we employed multistage cluster random sampling (*1–2*).

We first calculated the minimum required sample size for each stratum using Equation (1):

$$n_{min} = [z_{\alpha}^{2} p(1-p)/(p \times re)^{2}] \times deff \quad (1)$$

where z=1.96 (α =0.05), p=0.5 (environmental health literacy [EHL] prevalence assumed in the absence of prior data), re=0.15 (relative error), and $deff$=1.5 (design effect).

We then adjusted this minimum to determine the final target sample size Equation (2):

$$N = n_{\min} \times (\text{product of stratification factors}) \times (1 + \text{refusal rate}) \quad (2)$$

Stratification was based on gender (male, female) and survey area (urban, suburban), yielding four strata (2 × 2). Assuming a 14% refusal rate, the required sample size was at least 1,127 participants each year.

2020–2021

We targeted 3,720 interviews in 2020 and 3,100 in 2021 — representing three- and 2.5-fold multiples of the minimum required sample (1,127), respectively, with an additional 10% buffer to account for anticipated refusals. Ten Shanghai districts were selected through multistage cluster sampling: five urban districts (Xuhui, Changning, Huangpu, Putuo, and Yangpu) and five suburban districts (Songjiang, Qingpu, Fengxian, Jinshan, and Chongming). Within each district, administrative villages or neighborhood committees served as primary sampling units (PSUs), selected proportionally to population size. Each PSU randomly identified 75 households, from which one eligible resident aged 15–69 years was interviewed per household. Gender and age distributions were balanced to reflect local demographic characteristics. Each PSU was required to provide 60 valid questionnaires to meet the annual sampling targets.

2022–2024

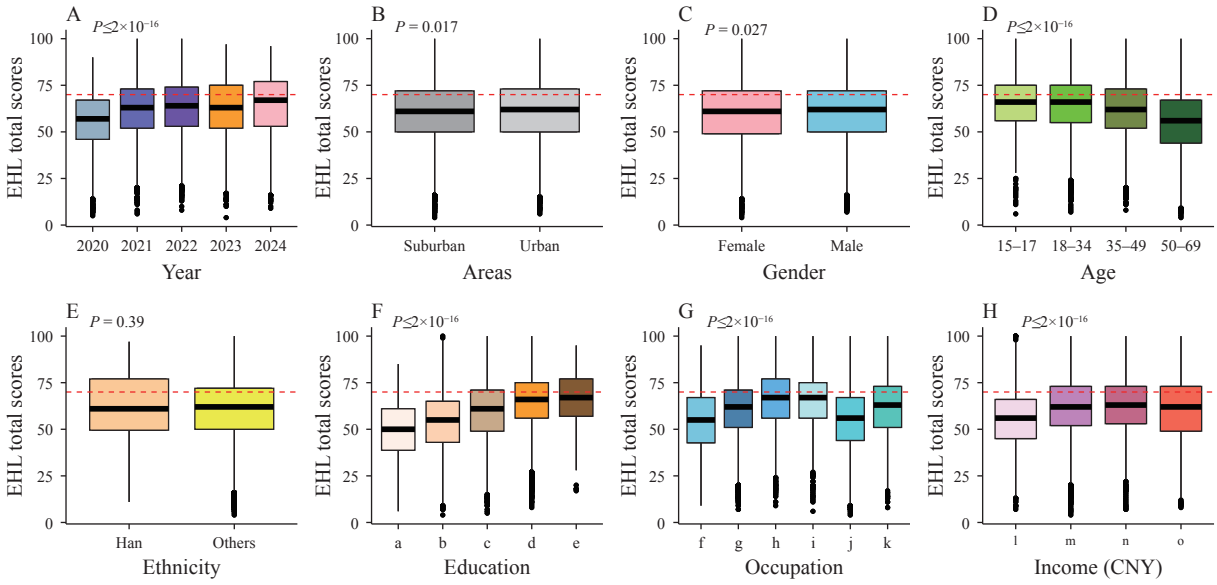
In accordance with the national ecological environment and health literacy monitoring protocol, we reduced coverage to six districts—three urban (Xuhui, Huangpu, and Yangpu) and three suburban (Qingpu, Fengxian, and Jinshan). Within each district, we selected four PSUs and implemented the same household sampling procedure: 75 households per PSU, one participant per household, with gender and age balanced to match local demographics. Each PSU yielded at least 60 valid questionnaires, ensuring a minimum of 1,440 participants annually.

Questionnaire Survey

Trained interviewers contacted eligible residents by telephone or home visits to obtain their consent and schedule appointments at either the participant's residence or the local Community Health Service Centre. Written informed consent was secured before each interview commenced. The survey protocol followed procedures established in previous studies (*3–5*).

The questionnaire consisted of two sections. Section A captured socio-demographic characteristics, including gender, age, education, and occupation. Section B incorporated the 47-item Core Questionnaire for Assessing the EHL of Chinese Residents, developed by the Ministry of Ecology and Environment (*1–2*, *3–5*). The assessment included 13 true/false questions (1 point each), 15 single-choice questions (2 points each), and 19 multiple-choice questions (3 points each). These items evaluated knowledge, attitudes, and behaviors regarding air quality, water safety, soil contamination, radiation exposure, noise pollution, ocean health, biodiversity conservation, climate change, household waste management, and toxic substance awareness. The total possible score was 100 points, with respondents scoring ≥ 70 classified as demonstrating adequate EHL (*1–2*, *3–5*).

Distribution of EHL Scores



SUPPLEMENTARY FIGURE S1. Distribution of total EHL scores across demographic subgroups.
Panels A–H display score distributions by survey year, geographic area, gender, age group, ethnicity, educational attainment, occupational category, and monthly income, respectively. Educational categories: a=primary school or below; b=junior high school; c=senior high school; d=vocational college/undergraduate; e=postgraduate and above. Occupational categories: f=farmer; g=city laborer; h=civil servant/leader; i=student; j=retiree; k=others. Income (CNY) categories l<5500, m=5500-12999, n=13000-20999, o≥21,000.
Abbreviation: EHL=Environmental Health Literacy, CNY=Chinese Yuan.

REFERENCES

1. Ministry of Ecological Environment of the People's Republic of China. Announcement on the release of the ecological environment and health literacy for Chinese citizens. 2020. https://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/202007/t20200727_791324.html. [2025-8-25]. (In Chinese).
2. Ministry of Ecology and Environment of the People's Republic of China. Technical guidelines for measuring citizens' environmental and health literacy (for trial implementation). 2017. https://www.mee.gov.cn/gkml/hbb/bgg/201706/W020170608498381052378.pdf. [2025-08-25]. (In Chinese).
3. Zhao Y, Sheng Y, Zhou JT, Wang H, Chilufya MM, Liu X, et al. Influencing factors of residents' environmental health literacy in Shaanxi province, China: a cross-sectional study. BMC Public Health 2022;22(1):114. https://doi.org/10.1186/s12889-022-12561-x.
4. Deng FD, Wen XY, Dong GQ, Wang XN, He HF, Zhu RX, et al. The environmental health literacy level was effectively improved of residents in Shaanxi Province, China, 2022. Front Public Health 2025;12:1499349. https://doi.org/10.3389/fpubh.2024.1499349.
5. Wang Q, Liu Q, Hou AL, Yu Y, Hu CH, Liang SW. Analysis on the current situation and influencing factors of environmental and health literacy of residents in Hubei province. Environ Sci Technol 2020;43(5):230 – 6. https://doi.org/10.19672/j.cnki.1003-6504.2020.05.031.

Preplanned Studies

Wastewater-based Surveillance of *Salmonella* Senftenberg as an Early-warning Indicator for Foodborne Outbreaks — Lianyungang City, Jiangsu Province, China, 2023–2025

Xiaolu Zhu[1]; Zhiyang Yao[1]; Jinli Huang[1]; Enbo Tao[1]; Jialing Zhang[1]; Haipeng Li[1]; Shengnan Cao[1]; Li Chen[1]; Huimin Qian[2,#]

Summary

What is already known about this topic?

Salmonella Senftenberg (*S.* Senftenberg) has emerged as a critical foodborne pathogen associated with major outbreaks. Currently, there is a lack of capacity for proactive risk prediction or preemptive containment of this serotype.

What is added by this report?

This study demonstrates that core-genome SNP analysis confirmed clonal identity between foodborne outbreak-associated and wastewater-derived ST14 S. Senftenberg isolates. Phylogenetically linked strains were identified in wastewater samples 7–14 days before the outbreak and persisted for more than 3 weeks.

What are the implications for public health practice?

Wastewater-based epidemiology (WBE) provides a critical early-warning signal for emerging foodborne outbreaks of serotype-specific *Salmonella*. When integrated with whole-genome sequencing (WGS), it offers distinct advantages in identifying cryptic transmission chains and undetected community-acquired foodborne infections.

ABSTRACT

Salmonella Senftenberg ST14 is recognized for its heightened environmental persistence and distinct virulence factors. It represents a significant public health threat through severe foodborne outbreaks and potential systemic infections such as septicemia. The current surveillance systems remain reactive as they lack proactive capabilities. This study demonstrates that core-genome single-nucleotide polymorphism analysis (≤6 differences) confirmed clonal identity between outbreak-associated food isolates and ST14 strains detected in wastewater. Phylogenetically linked strains were identified in wastewater samples 7–14 days before the outbreak and persisted for more than 3 weeks. Subsequent 6-month wastewater surveillance confirmed the sustained community circulation of ST14, corroborated by environmental contamination at the implicated facilities. A 24-month monitoring demonstrated no outbreak recurrence, indicating the absence of outbreak-associated strains in the wastewater. These findings suggest that wastewater-based surveillance can serve as an early warning of emerging serotype-specific *Salmonella* outbreaks and in tracking transmission dynamics at the population level. Furthermore, whole-genome sequencing (WGS)-enhanced microevolutionary analysis provides distinct advantages for uncovering cryptic transmission chains and identifying undetected community-acquired foodborne infections, thereby enabling more targeted public health interventions.

Wastewater-based Epidemiology (WBE) initially focused on monitoring chemical analytes (*1*), such as tobacco constituents, illicit substances, pharmaceuticals, and metabolites. Its scope has progressively expanded to microbial surveillance through technological advances. Notably, these include human pathogens such as *Salmonella enterica*, *Campylobacter jejuni* (*2*), diarrheagenic *Escherichia coli* (*3*), SARS-CoV-2 (*4*), and hepatitis A virus (*5*). SARS-CoV-2 sewage surveillance has demonstrated advantages as an early warning of clinically reported variants (*4*). In contrast, no peer-reviewed studies have yet documented the detection of pathogenic bacteria in sewage before clinical reports.

Salmonella Senftenberg (*S.* Senftenberg) has emerged as a critical foodborne pathogen linked to major outbreaks, owing to its enhanced environmental resilience and distinct virulence factors. Severe *S.* Senftenberg infections can progress to life-threatening systemic diseases such as

septicemia. However, the current food safety surveillance systems remain retrospective and lack the capacity for proactive risk prediction or preemptive containment of this serotype.

Successive foodborne disease outbreaks caused by *S.* Senftenberg occurred between May 28 and June 7, 2023, in Donghai County, Lianyungang City, China. Concurrently, *S.* Senftenberg strains were isolated from three consecutive sewage samples from the Haizhou District, and all isolates exhibited indistinguishable pulsed-field gel electrophoresis (PFGE) patterns. This retrospective study assessed the early-warning capacity of sewage-based genomic analysis by examining clinical outbreaks and environmental verification strains from Donghai County, sewage isolates from Haizhou and Lianyun districts, and *S.* Senftenberg isolates archived in China's Pathogen Identification Network (PIN). This study integrated genomic data from clinical outbreaks, environmental surveys, and a prospective wastewater surveillance program to evaluate the potential of WBE for the early warning and source tracking of *S.* Senftenberg outbreaks.

This retrospective study was conducted in Lianyungang City, Jiangsu Province, China, after consecutive foodborne disease outbreaks caused by *S.* Senftenberg in Donghai County (May 28–June 7, 2023). It integrated data from clinical outbreaks, environmental verification surveys, a 2-year prospective sewage surveillance study (April 18, 2023, to April 17, 2025), and China's PIN (2015–2025).

A total of 73 *S.* Senftenberg isolates were analyzed, comprising 20 clinical and environmental isolates from the 2023 Donghai County outbreaks, 42 sewage-derived isolates from prospective surveillance, 8 isolates from environmental verification surveys, and 3 historical clinical isolates from Lianyungang obtained from China's PIN.

Salmonella strains were isolated and identified according to the methods described by Sun et al (*6*). Presumptive *Salmonella* colonies were serotyped using the Kauffmann–White scheme. PFGE was performed on all outbreak-associated isolates (n=20) and all *S.* Senftenberg isolates obtained from sewage during the outbreak period (n=4) using *Xba*I restriction. The resulting patterns were analyzed with BioNumerics software (version 7.6, Applied Maths NV, Sint-Martens-Latem, Belgium).

Genomic DNA was extracted from all the isolates using a FastPure® Bacteria DNA Isolation Mini Kit (Vazyme, Nanjing, China). WGS was performed on an Illumina NovaSeq platform (Illumina, San Diego, CA, USA) using the 150 bp paired-end mode. Sequence types were determined using CLC Genomics Workbench. Core-genome single-nucleotide polymorphism (SNP) calling was performed using Snippy, and recombination removal was performed using Gubbins. A core-genome SNP-based phylogenetic tree was constructed using RAxML software (version 8.2.12; Alexandros Stamatakis, Heidelberg Institute for Theoretical Studies, Heidelberg, Germany).

Each batch included negative and positive controls (*Salmonella* Typhimurium ATCC 14028). The culture media and critical equipment were subjected to regular quality checks and calibration. All data were logged using a Laboratory Information Management System for traceability. Molecular characterization was performed using PFGE following a standardized protocol with strain H9812.

Spearman's rank correlation coefficient was calculated using the Statistical Package for the Social Sciences software (version 25.0, IBM Corp., Armonk, NY, USA). Spearman's rho (r_s) and its exact p values are reported. A two-tailed P<0.05 was considered statistically significant.

Four distinct foodborne disease outbreaks caused by *S.* Senftenberg occurred between May 28 and June 7, 2023. The outbreak series began with two events associated with banquets at Restaurant E: Event 1, May 28, 2023; and Event 2, June 2, 2023. Subsequently, a secondary transmission cluster (Event 3) emerged on June 2, involving four family members infected through contaminated marinated pig ears purchased from Supermarket A. The final outbreak (Event 4) occurred at Restaurant B on June 5 (Figure 1). Epidemiological tracing identified Trading Company C as the common distribution hub for Supermarket A and Restaurant B, and further traceback investigations implicated Food Processing Plant D as the contamination source. Among the 79 specimens tested, 20 *S.* Senftenberg isolates were recovered from seven clinical cases, five asymptomatic food handlers, seven food specimens, and one environmental specimen (Supplementary Table S1, available at https://weekly.chinacdc.cn/).

A 2-year prospective sewage surveillance study was conducted in Lianyungang from April 18, 2023, to April 17, 2025. Samples of *Salmonella* spp. were collected weekly from the Dapu Sewage Treatment Plant (STP) in the Haizhou District from April to December 2023. Sampling was expanded to the Xugou STP in the Lianyun District, beginning in January

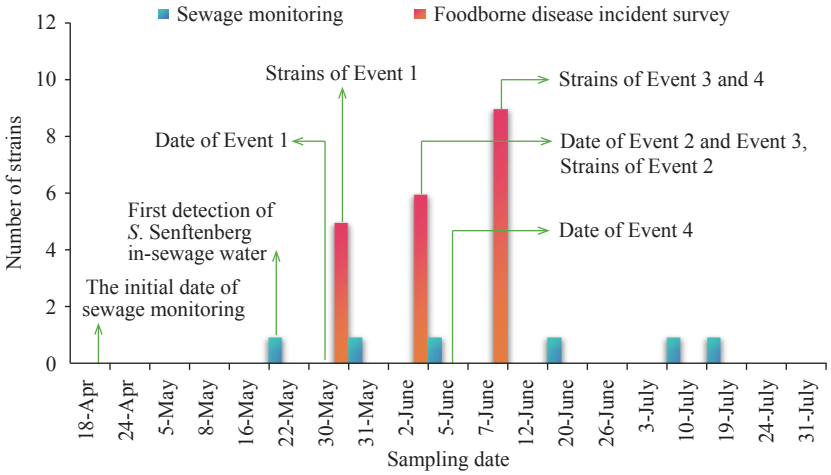


FIGURE 1. Temporal distribution of *Salmonella* Senftenberg isolated from sewage monitoring and foodborne outbreaks over time in Lianyungang, April–July 2023.
Note: The epidemiological curve indicates that sewage monitoring began on April 18, 2023. The first *S.* Senftenberg strain was isolated from sewage on May 22, 2023, 1 week before Event 1. The strain was continuously detected until June 5, 2023, with three isolated strains (indicated by blue pillars). A separate progression of *S.* Senftenberg strains isolated from foodborne disease cases began on May 28, 2023, with Event 1 continuing through Event 4 (indicated by red pillars). All detailed sewage monitoring and foodborne disease incidents are marked on the corresponding dates.

2024. Haizhou and Lianyun Districts lie east of Donghai County, where the outbreak was identified; Haizhou District is geographically adjacent to Donghai County. From 171 samples, 42 *S.* Senftenberg isolates were recovered. The ST14 subtype was predominant (*n*=34, 80.95%), followed by minor subtypes including ST185 (*n*=6, 14.29%), ST217 (*n*=1, 2.38%), and ST684 (*n*=1, 2.38%). Five rounds of sewage surveillance were conducted during the pre-outbreak phase (April 18–May 21). A total of 22 *Salmonella* strains were isolated, including *S.* Agona (*n*=4), *S.* Kentucky (*n*=3), *S.* Muenster (*n*=3), *S.* London (*n*=3), *S.* Bovismorbificans (n=2), *S.* Derby (*n*=1), *S.* Enteritidis (*n*=1), *S.* Reading (*n*=1), *S.* Sandow (*n*=1), *S.* Infantis (*n*=1), *S.* Corvallis (*n*=1), and *S.* Mississippi (*n*=1). However, no serotypic strains associated with subsequent outbreaks were identified. The first sewage-derived *S.* Senftenberg isolate was detected on May 22, 2023, 7 days before the onset of the initial outbreak. During the outbreak, all three strains were sequentially recovered. Sewage isolates of *S.* Senftenberg demonstrated 100% PFGE pattern congruence with the clinical outbreak strains (Figure 2). Notably, an ST14 variant from a sample collected on June 20, 2023, exhibited a divergent PFGE pattern, with 91% similarity to outbreak strains, indicating post-outbreak genomic diversification. The association between the weekly detection of *S.* Senftenberg in wastewater and the number of reported foodborne disease outbreak isolates in the subsequent 7 days was analyzed. The analysis revealed a statistically significant and strong positive correlation [$r_s(8)=0.759$, $P=0.011$]. These findings provide statistical evidence supporting the effectiveness of wastewater surveillance as an early-warning system for *Salmonella* outbreaks.

Following the persistent detection of *S.* Senftenberg in sewage after the outbreak resolution, a targeted survey was conducted on November 28, 2023, at the facilities implicated in Donghai County to evaluate the risk of recurrence. A total of 68 samples were collected, including food products, raw ingredients, food handler specimens, environmental swabs, and drainage water. Eight *S.* Senftenberg ST14 strains were isolated from several sites, including four from Food Processing Plant D, two from Supermarket A, and two from Restaurant E (Supplementary Table S2, available athttps://weekly.chinacdc.cn/). Rigorous disinfection protocols were implemented on all affected premises.

From 2015 to 2025, five *S.* Senftenberg isolates were obtained from Lianyungang via the China's PIN. These isolates included one ST14 strain from 2017 from a 15-year-old patient with diarrhea, two ST14 strains from 2022 from infants aged 6 months and 1 year with diarrhea, one ST1959 strain associated with co-infection in 2024, and one ST185 strain recovered from contaminated chili powder in 2024.

A Bayesian phylogenomic analysis of core-genome SNPs revealed ≤6 SNP differences among the three epidemiologically linked groups: 20 clinical isolates from the 2023 Donghai County foodborne outbreak,

Isolate ID	Collection date	Source	Event	Place
SM_E027	2023-06-05	Sewage	–	Haizhou
SM_D064	2023-05-22	Sewage	–	Haizhou
SM_D073	2023-05-31	Sewage	–	Haizhou
SM_E017	2023-06-07	Human faeces	4	Donghai
SM_E018	2023-06-07	Pork ear	3	Donghai
SM_E019	2023-06-07	Pork ear	3	Donghai
SM_E020	2023-06-07	Human faeces	4	Donghai
SM_E021	2023-06-07	Human faeces	4	Donghai
SM_E022	2023-06-07	Human faeces	3,4	Donghai
SM_E023	2023-06-07	Food container	3,4	Donghai
SM_E024	2023-06-07	Pork ear	4	Donghai
SM_E025	2023-06-07	Duck gizzard	4	Donghai
SM_E006	2023-05-30	Human faeces	1	Donghai
SM_E007	2023-05-30	Human faeces	1	Donghai
SM_E008	2023-05-30	Human faeces	1	Donghai
SM_E009	2023-05-30	Human faeces	1	Donghai
SM_E010	2023-05-30	Human faeces	1	Donghai
SM_E011	2023-06-02	Pork mixture	2	Donghai
SM_E012	2023-06-02	Pork ear	2	Donghai
SM_E013	2023-06-02	Preserved egg	2	Donghai
SM_E014	2023-06-02	Human faeces	2	Donghai
SM_E015	2023-06-02	Human faeces	2	Donghai
SM_E016	2023-06-02	Human faeces	2	Donghai
SM_E038	2023-06-20	Sewage	–	Haizhou

FIGURE 2. Dendrogram of PFGE patterns of 24 *Salmonella* Senftenberg isolates obtained from sewage monitoring and foodborne disease outbreak surveillance in the same period.
Note: "–" indicates no declared association with the events.
Abbreviation: PFGE=pulsed-field gel electrophoresis.

3 contemporaneous wastewater isolates from the Haizhou District, and 6 post-outbreak environmental verification strains. These isolates established a clonal outbreak cluster, demonstrating an evolutionary divergence from other clinical and environmental strains collected between 2023 and 2025 (Figure 3). The remaining 31 sewage-derived isolates exhibited 0–134 SNP differences. The high genetic diversity among sewage isolates suggests a complex ecology within the wastewater system, either promoting microevolution or indicating the introduction of multiple distinct lineages over time. For example, two strains from the post-outbreak environmental verification were genetically identical (without SNP differences), as were those from the 2022 diarrheal cases. In contrast, the isolate from the 2017 diarrheal case displayed discernible SNP variations.

DISCUSSION

This study conducted a 2-year longitudinal surveillance of *S.* Senftenberg in the urban wastewater systems of Lianyungang City. Genomic analysis revealed that all ST14 isolates from foodborne outbreaks, concurrently sampled urban wastewater, and most post-outbreak environmental verification surveys formed a single outbreak clone. Notably, wastewater isolates were detected 1 week before the emergence of clinical cases. These findings highlight the potential of urban wastewater surveillance as an early warning system for foodborne disease outbreaks, enabling population-level tracking of transmission dynamics and the detection of cryptic transmission through high-resolution genomic profiling.

This study statistically validates the utility of wastewater monitoring as an early warning system. The strong correlation [$r_s(8)=0.759$] provides quantitative evidence, solidifying the scientific case for its adoption by public health authorities for proactive outbreak surveillance. The *Salmonella* Java variant detected in the United States wastewater emerged after clinical outbreaks linked to imported tuna (*7*), and ST14 *S.* Senftenberg in Lianyungang wastewater was detected

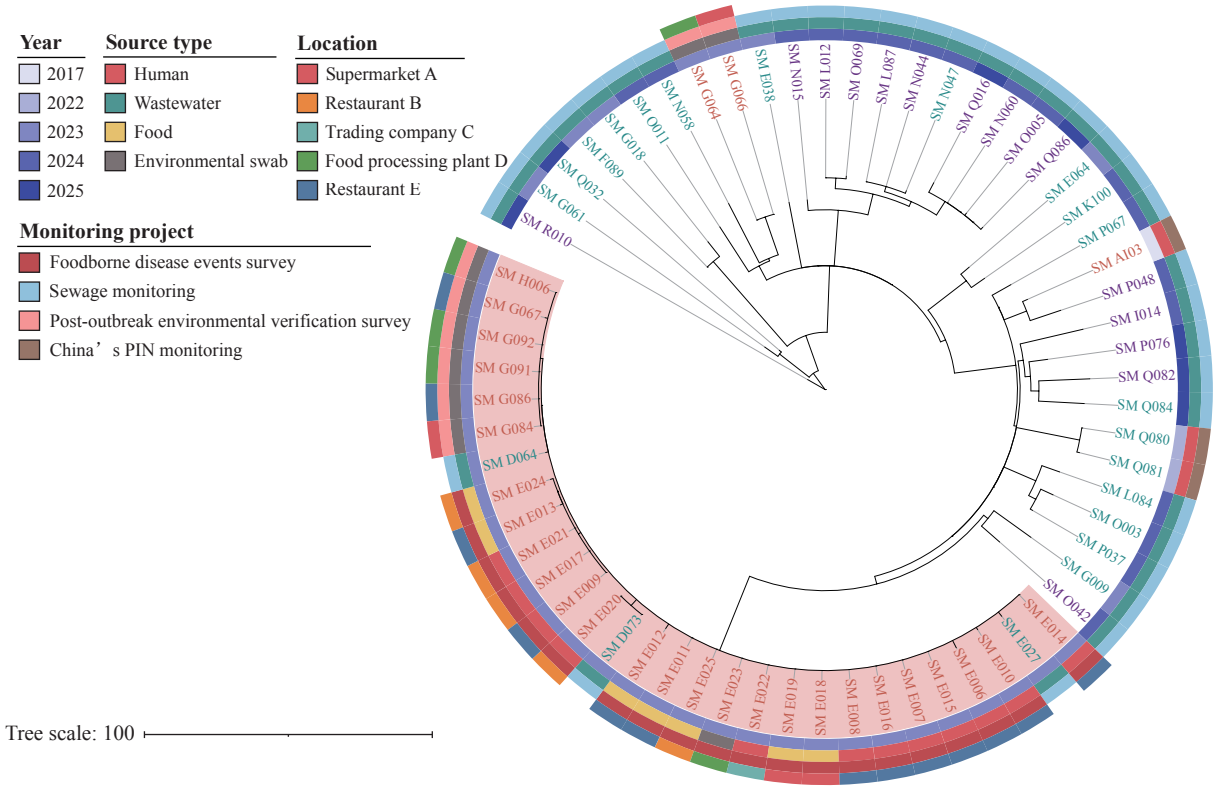


FIGURE 3. Phylogenetic trees of 65 *Salmonella* Senftenberg ST14 strains.
Note: From the inner to the outer circle, the rings indicate: (1) year of isolation; (2) source type; (3) monitoring project of origin; (4) geographic location of incident-related strains. Label colors denote specific locations: brown represents strains isolated from Donghai County, green represents strains isolated from Haizhou District, and purple represents strains isolated from Lianyun District. The pink-highlighted clade represents the phylogenetic branch containing outbreak-associated strains.
Abbreviation: SNP=single-nucleotide polymorphism.

on May 22, 7 days before the foodborne incident reported on May 29. *S.* Senftenberg was undetectable in five wastewater samples collected between April 18 and May 21. However, it was consistently detected from May 22 onward, suggesting *de novo* pathogen introduction into the wastewater system, rather than residual shedding. Potential routes of introduction include influxes from external reservoirs (*8*) or undetected human sources. However, this study lacked direct evidence to differentiate these pathways, leaving the underlying mechanism unclear. In future sewage surveillance, consistently detecting a previously undetected *Salmonella* serotype with uniform molecular banding patterns may signal potential cryptic transmission or an elevated risk of an impending foodborne outbreak. Enhanced surveillance through increased sampling frequency and sample size is recommended, with immediate source-tracking investigations to identify potential contamination sources.

WBE is a critical complementary method for characterizing outbreak strain dissemination and monitoring epidemic progression; this is particularly relevant for non-typhoidal *Salmonella* (NTS) because high rates of subclinical infection often lead to an underestimated prevalence (*9*). In Lianyungang, longitudinal surveillance revealed the sustained detection of *S.* Senftenberg. The isolation percentages were 32.43% in 2023 and 18.27% in 2024, which exceeded the corresponding clinical reporting rates and suggesting potential environmental persistence. Furthermore, ST14 strains of *S.* Senftenberg were identified in Haizhou wastewater 6 months following the outbreak, indicating either extended environmental survival or ongoing low-level transmission. These findings prompted a systematic re-investigation of the primary outbreak sites, verifying that residual environmental persistence is linked to suboptimal decontamination. After a 24-month surveillance period, the strain disappeared from the wastewater system, confirming the end of cryptic community transmission. The WBE enables population-scale

burden estimation and transmission chain reconstruction of diverse NTS serovars (*Salmonella* Java, Derby, and Senftenberg) (*7–9*), facilitating evidence-based intervention strategies.

WGS exhibits an exceptional discriminatory power in wastewater surveillance, serving as an unparalleled tool for identifying outbreak strains and profiling microbial diversity. For example, consecutive biennial wastewater monitoring in Lianyungang revealed extensive genetic variation (0–134 core SNPs) in *S.* Senftenberg populations, demonstrating the hallmarks of adaptive evolutionary capacity within complex aquatic ecosystems. Wastewater environments are highly dynamic, with varied selective pressures that may shape the bacterial genomic architecture. In sewage monitoring, whole-genome SNP phylogenetics has been used to resolve strain-specific evolutionary trajectories. This method can phylogenetically discriminate between outbreak, sporadic, and environmental clusters while identifying the specific clone responsible for foodborne outbreaks. This high-resolution method is in stark contrast to PFGE profiling. PFGE is limited by its phylogenetic resolution, which indicates <100% pattern similarity, and cannot clearly delineate outbreak clusters. The absence of outbreak-associated strains in the wastewater suggests the termination of community transmission following the intervention. However, the specific evolutionary mechanisms responsible for strain clearance could not be determined from the available data.

When wastewater surveillance identifies pathogens with potential foodborne outbreaks, an immediate multi-pronged response is required; this includes enhanced monitoring of major commercially available foods and drinking water across the sewage catchment area to enable early identification of transmission sources. Active screening should be initiated for wastewater from relevant communities and among diarrheal cases in healthcare settings to accurately identify infected individuals and define the outbreak's scope. Furthermore, wastewater treatment processes should be regularly reviewed to ensure compliance with health standards and to minimize secondary environmental contamination.

Several advanced methodologies have been recommended to trace the origins of these high-risk strains. First, reverse tracing can be conducted through the sewage network using methods such as catchment area investigation and integrating environmental and epidemiological investigations, including infectious case contact tracing and inspection of high-risk premises. Advanced technologies, such as microbial source tracking and WGS, can also be applied. Implementing these measures creates a robust routine monitoring, and early warning system. By establishing wastewater surveillance as a sentinel in the public health network, potential threats can be detected earlier, leading to a more effective community health protection.

This study had two limitations. First, Donghai's outbreak epicenter and Haizhou's wastewater surveillance sites are separated by approximately 40km and lack a shared sewer shed connection. Therefore, whether the wastewater signal originates from a different undetected outbreak contributing to the wastewater plant remains undetermined. Second, although the transmission chain was partially traced, the initial source remains unknown, such as the original contaminated food product entering the processing plant.

Integrated WBE-WGS platforms provide high-resolution spatiotemporal pathogen tracking and generate operational intelligence for predictive outbreak modeling. Compared to conventional clinical systems, multimodal surveillance enables earlier detection of epidemic phases by approximately 7–14 days, empirically validating its enhanced early warning systems.

Conflicts of interest: No conflicts of interest.

Ethical statement: Approval by the Declaration of Helsinki and approved by the Ethics Committee of the Jiangsu Provincial Center for Disease Prevention and Control (JSJK2024-B011-01).

Funding: Supported by the National Key R&D Program of China (2024YFC2311302) and the Preventive Medicine Research Project of Jiangsu Province (Nos. Ym2023040 and Ym2023082).

doi: 10.46234/ccdcw2025.268

[#] Corresponding author: Huimin Qian, jsqhm@jscdc.cn.

[1] Lianyungang Center for Disease Control and Prevention, Lianyungang City, Jiangsu Province, China; [2] NHC Key Laboratory of Enteric Pathogenic Microbiology, Jiangsu Provincial Center for Disease Prevention and Control, Nanjing City, Jiangsu Province, China.

Submitted: August 20, 2025
Accepted: November 24, 2025
Issued: December 19, 2025

REFERENCES

1. Castiglioni S, Senta I, Borsotti A, Davoli E, Zuccato E. A novel approach for monitoring tobacco use in local communities by wastewater analysis. Tob Control 2015;24(1):38 – 42. https://doi.org/10.1136/tobaccocontrol-2014-051553.
2. Bonetta S, Pignata C, Lorenzi E, De Ceglia M, Meucci L, Bonetta S, et al. Detection of pathogenic *Campylobacter, E. coli* O157:H7 and Salmonella spp. in wastewater by PCR assay. Environ Sci Pollut Res Int 2016;23(15):15302 – 9. https://doi.org/10.1007/s11356-016-6682-5.
3. Yang K, Pagaling E, Yan T. Estimating the prevalence of potential enteropathogenic *Escherichia coli* and intimin gene diversity in a human community by monitoring sanitary sewage. Appl Environ Microbiol 2014;80(1):119 – 127. https://doi.org/10.1128/AEM.02747-13.
4. Wongprommoon A, Chomkatekaew C, Chewapreecha C. Monitoring pathogens in wastewater. Nat Rev Microbiol 2024;22(5):261. https://doi.org/10.1038/s41579-024-01033-1.
5. Braunfeld JB, Dao BL, Buendia J, Amiling R, LeBlanc C, Jewell MP, et al. *Notes from the Field*: genomic and wastewater surveillance data to guide a hepatitis a outbreak response — los Angeles county, March 2024-June 2024. MMWR Morb Mortal Wkly Rep 2025;74(5):66 – 8. https://doi.org/10.15585/mmwr.mm7405a3.
6. Sun HH, Ling XM, Li Y, Li Y, Cui SH, Bai L. Research on quantitative method and contamination level of *Salmonella enterica* in raw pork from farmer's markets in Chengdu. Chin J Prev Med 2021;55(8):999 – 1005. https://doi.org/10.3760/cma.j.cn112150-20210302-00209.
7. Diemert S, Yan T. Clinically unreported salmonellosis outbreak detected via comparative genomic analysis of municipal wastewater *Salmonella* isolates. Appl Environ Microbiol 2019;85(10):e00139 – 19. https://doi.org/10.1128/AEM.00139-19.
8. Naudin SA, Ferran AA, Imazaki PH, Arpaillange N, Marcuzzo C, Vienne M, et al. Development of an *in vitro* biofilm model for the study of the impact of fluoroquinolones on sewer biofilm microbiota. Front Microbiol 2024;15:1377047. https://doi.org/10.3389/fmicb.2024.1377047.
9. Diemert S, Yan T. Municipal wastewater surveillance revealed a high community disease burden of a rarely reported and possibly subclinical *Salmonella enterica* serovar derby strain. Appl Environ Microbiol 2020;86(17):e00814 – 20. https://doi.org/10.1128/AEM.00814-20.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. *Salmonella* Senftenberg strains isolated from foodborne disease events survey.

Isolate ID	ST	Collection date	Source	Incident number	Location
SM_E006	ST14	2023/5/30	Human feces	Event 1	Restaurant E
SM_E007	ST14	2023/5/30	Human feces	Event 1	Restaurant E
SM_E008	ST14	2023/5/30	Human feces	Event 1	Restaurant E
SM_E009	ST14	2023/5/30	Human feces	Event 1	Restaurant E
SM_E010	ST14	2023/5/30	Human feces	Event 1	Restaurant E
SM_E011	ST14	2023/6/2	Pork mixture	Event 2	Restaurant E
SM_E012	ST14	2023/6/2	Pork ear	Event 2	Restaurant E
SM_E013	ST14	2023/6/2	Preserved egg	Event 2	Restaurant E
SM_E014	ST14	2023/6/2	Human feces	Event 2	Restaurant E
SM_E015	ST14	2023/6/2	Human feces	Event 2	Restaurant E
SM_E016	ST14	2023/6/2	Human feces	Event 2	Restaurant E
SM_E017	ST14	2023/6/7	Human feces	Event 4	Restaurant B
SM_E018	ST14	2023/6/7	Pork ear	Event 3	Supermarket A
SM_E019	ST14	2023/6/7	Pork ear	Event 3	Supermarket A
SM_E020	ST14	2023/6/7	Human feces	Event 4	Restaurant B
SM_E021	ST14	2023/6/7	Human feces	Event 4	Restaurant B
SM_E022	ST14	2023/6/7	Human feces	Event 3,4	Trading Company C
SM_E023	ST14	2023/6/7	Food container	Event 3,4	Processing Plant D
SM_E024	ST14	2023/6/7	Pork ear	Event 4	Restaurant B
SM_E025	ST14	2023/6/7	Duck gizzard	Event 4	Restaurant B

SUPPLEMENTARY TABLE S2. *Salmonella* Senftenberg strains isolated from the post-outbreak environmental verification survey.

Isolate ID	ST	Collection date	Source	Location
SM_G064	ST14	2023/11/28	Ground	Processing Plant D
SM_G066	ST14	2023/11/28	Chopping board (cooked food)	Supermarket A
SM_G067	ST14	2023/11/28	Kitchen sewer sewage	Restaurant E
SM_G084	ST14	2023/11/28	Chopping board (raw food)	Supermarket A
SM_G086	ST14	2023/11/28	Chopping board and knife	Restaurant E
SM_G091	ST14	2023/11/28	Toilet	Processing Plant D
SM_G092	ST14	2023/11/28	Brine	Processing Plant D
SM_H006	ST14	2023/11/28	Food container	Processing Plant D

Preplanned Studies

A Multi-omics Framework Combining Genomics and Proteomics for Silicosis Prediction in Chinese Workers — Jiangsu Province, China, 2023–2024

Furu Wang[1,2]; Qianqian Gao[2]; Chenjie Li[1]; Lei Han[2]; Chuanfeng Zhang[2]; Zhengdong Zhang[1,#]; Baoli Zhu[2,#]

Summary

What is already known about this topic?

Currently, the detection of silicosis relies on imaging and pulmonary function tests, which are effective only at identifying the advanced stages. Additionally, no effective protein biomarkers or genetic risk models exist for the early detection or targeted intervention of silicosis.

What is added by this report?

This study integrates genomics and proteomics to identify new genetic loci associated with susceptibility to silicosis. Using Mendelian randomization and protein quantitative trait loci (pQTL) analysis, 2 functionally significant genetic variants [rs6677666 (WLS) and rs2272528 (COL4A4)] and 5 protein biomarkers (MMP12, EGF, Gal_9, GZMA, and ICOSLG) mechanistically linked to silicosis pathogenesis were identified. A diagnostic causal protein risk score (CPRS) model was then constructed to provide a robust tool for early detection in high-risk populations.

What are the implications for public health practice?

These findings provide new insights into the early diagnosis of silicosis, and support the development of preventive and screening strategies for populations at risk, enhancing public health policies for the control and management of silicosis.

ABSTRACT

Objective: In this study, we aimed to identify novel genetic loci and protein biomarkers associated with silicosis susceptibility in Chinese workers through integrated proteomic and genomic analyses and to develop an early diagnostic prediction model.

Methods: A genome-wide association study (GWAS) was conducted on 163 patients with silicosis and 183 controls, followed by Olink proteomic profiling of 92 plasma proteins. Protein quantitative trait loci (pQTL) mapping, Mendelian randomization (MR), and Bayesian co-localization were used to infer causal relationships. A causal protein risk score (CPRS) model integrating genetic and proteomic data was developed and validated using 10-fold cross-validation.

Results: GWAS identified 16 novel risk loci ($P<1\times10^{-5}$), including rs6677666 (WLS) and rs2272528 (COL4A4). MR analysis revealed eight plasma proteins associated with silicosis risk, with MMP12, EGF, Gal_9, GZMA, and ICOSLG showing significant differential expression ($P<0.05$). The CPRS model combining these proteins demonstrated a high diagnostic accuracy (AUC=0.915), outperforming traditional clinical variables.

Conclusion: This multi-omics study uncovered genetic and proteomic markers linked to silicosis susceptibility and established a robust predictive model. The integration of GWAS and proteomics offers novel insights into the pathogenesis of silicosis, and supports development of early detection and prevention policies for high-risk populations.

Silicosis is an occupational lung disease caused by the inhalation of respirable crystalline silica (RCS) dust and is characterized by chronic inflammation, collagen deposition, and fibrotic lesions that drive a continuous progression from subclinical pathology to severe lung tissue damage (*1*). Currently, no effective cure exists (*2*). Early detection and timely intervention are critical strategies for extending the life expectancy of individuals with silicosis. Recent advances in high-throughput technologies and bioinformatics have enabled comprehensive multi-omics analyses. Therefore, in this study, we aimed to identify novel protein biomarkers and potential therapeutic targets for silicosis to inform public health strategies for its prevention and treatment. First, we conducted a

genome-wide association study (GWAS) involving 163 patients with silicosis and 183 controls, followed by Olink proteomic profiling of 92 plasma proteins. Protein quantitative trait loci (pQTL) analysis, Mendelian randomization (MR), and Bayesian co-localization were used to infer causal relationships. Finally, a causal protein–based risk score, incorporating proteins showing consistent causal evidence, was developed and evaluated for its association with silicosis risk and predictive performance using appropriate statistical models and validation techniques.

Study participants included patients with silicosis undergoing medical follow-up and retired workers with a history of silica dust exposure but without a diagnosis of pneumoconiosis (control group). The control group was matched to the silicosis group based on the duration of silica dust exposure. Data collection included demographics (sex and age), clinical parameters (pneumoconiosis type), occupational factors (occupation type and cumulative dust exposure), lifestyle factors (smoking and drinking habits), and workplace safety information (personal protective equipment usage). The silicosis classification followed the International Labour Organization (ILO) guidelines (*3*).

GWAS was conducted using 163 patients and 183 controls selected from the aforementioned participants. Genotyping was performed using Illumina Infinium Asian Screening Array-24 v1.0 BeadChip platform (Illumina Inc., San Diego, CA, USA), which was specifically designed for East Asian populations. This chip contains approximately 670,000 markers and provides robust coverage of low-frequency variants [minor allele frequency (MAF)=1%–5%] common in the East Asian region. The baseline characteristics of the participants are provided in the Supplementary Table S1 (available at https://weekly.chinacdc.cn/).

Following quality control, 157 patients with silicosis and 182 controls were genotyped. A total of 7,897,033 autosomal single-nucleotide polymorphisms (SNPs) passed the quality control standards and were included for further analysis. Manhattan and quantile–quantile plots for the GWAS results are presented in Supplementary Figure S1 (available at https://weekly.chinacdc.cn/). After adjusting for potential confounding factors, including age, sex, dust concentration, and pack-years of smoking, along with the first 10 principal components, we identified 42,210 risk loci with $P<0.05$. Among these, 16 independent risk loci achieved genome-wide significance ($P<1\times10^{-5}$): (Table 1 and Supplementary Figure S2, available at https://weekly.chinacdc.cn/). However, these loci have not been previously reported to be associated with silicosis.

Among these, rs3738756, rs28536654, rs4776983,

TABLE 1. Sixteen newly identified risk loci for silicosis in the Chinese population.

UniqID	rsID	chr	pos	Allele A	Allele B	Annotation	*OR*	*P* (1×10^{-5})	nSNPs	Nearest gene
1:68640101:C:T	rs6677666	1	68640101	C	T	Intronic	0.373	0.505	25	WLS
1:110658260:C:T	rs3738756	1	110658260	C	T	Intergenic	2.555	0.339	32	UBL4B
2:227976749:C:G	rs2272528	2	227976749	C	G	Intronic	0.378	0.316	22	COL4A4
4:57963611:A:G	rs4865182	4	57963611	A	G	Intronic	0.417	0.695	4	IGFBP7
5:26013029:A:G	rs6885607	5	26013029	A	G	Exonic	0.109	0.709	30	RNU4-43P
5:26324525:A:G	rs112721803	5	26324525	A	G	Intronic	0.136	0.365	104	RP11-351N6.1
5:164319653:G:T	rs17070163	5	164319653	G	T	Intronic	0.418	0.464	11	LINC03000
6:112154244:A:T	rs1391373	6	112154244	A	T	Intergenic	0.314	0.872	93	HLA-DQA1
7:34543087:C:T	rs7802435	7	34543087	C	T	Intronic	0.427	0.877	4	NPSR1-AS1
10:47737036:C:T	rs28536654	10	47737036	C	T	Upstream	2.250	0.848	1	Lnc-ANXA8L1-1
12:113444117:C:G	rs2240185	12	113444117	C	G	Intronic	0.427	0.666	10	OAS2
15:68115177:A:G	rs4776983	15	68115177	A	G	Intronic	5.882	0.072	1	SKOR1
16:83103233:C:T	rs9923246	16	83103233	C	T	Intronic	2.653	0.101	5	CDH13
16:86383556:A:G	rs7191547	16	86383556	A	G	Intronic	0.368	0.769	3	LINC00917
18:35461186:A:G	rs28814664	18	35461186	A	G	Intergenic	0.307	0.965	42	MIR4318
21:31135702:C:T	rs72551940	21	31135702	C	T	Exonic	0.339	0.774	2	GRIK1

Abbreviation: SNPs=single-nucleotide polymorphisms; *OR*=odds ratio.

and rs9923246 were positively correlated ($P<1\times10^{-5}$) with the risk of developing silicosis. Gene expression regulation analysis revealed that the G mutation of rs4865182 downregulated *IGFBP7* gene expression in the blood. *IGFBP7* is a detrimental factor that exacerbates acute inflammatory responses such as acute lung injury (*4*).

Genome-wide pQTL analysis of 182 healthy controls identified 689 significant cis-pQTL variant–protein interactions ($P<0.05$). In addition to transcriptional regulatory regions, such as promoters and enhancers, pQTL variants were also found in post-transcriptional regulatory regions, such as untranslated regions (UTRs). The cis-pQTL variants were predominantly located in the intronic regions of adjacent genes (51%), suggesting that introns play a significant role in the regulation of protein expression.

Proteomic MR analysis identified eight proteins that were significantly correlated with silicosis ($P<0.05$). Four proteins showed positive associations: MMP12 [odds ratio (*OR*)=4.717, $P=0.027$], CD244 (*OR*=3.251, $P=0.039$), GZMA (*OR*=7.520, $P=0.002$), and ICOSLG (*OR*=12.351, $P=0.016$); four proteins demonstrated negative effects: EGF (*OR*=0.129, $P=0.014$), MUC_16 (*OR*=0.442, $P=0.005$), Gal_9 (*OR*=0.022, $P=0.003$), and CD28 (*OR*=0.367, $P=0.012$) (Figure 1).

To assess the causal relation between proteins and different stages of silicosis, we performed a GWAS

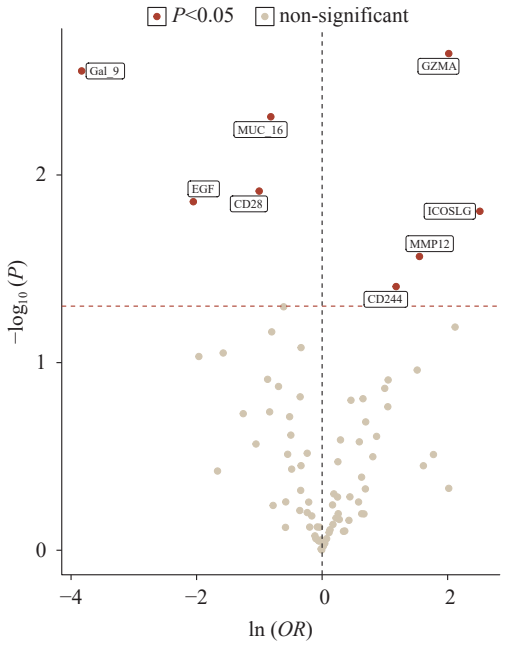


FIGURE 1. Protein-wide MR of cis-pQTLs and silicosis.
Abbreviation: MR=mendelian randomization.

stratified by disease stage. A total of 134 patients with stage I disease vs. healthy controls and 6,172,798 variants were included in the subsequent analysis. MR and pQTL analyses identified CD28, GZMA, and MMP12 as causally associated with Stage I disease. In an analysis of 22 patients with stage II disease *vs.* healthy controls, only CD244 exhibited a causal association.

Among the eight proteins with causal associations, we selected five (MMP12, EGF, Gal_9, GZMA, and ICOSLG) that demonstrated significant differential expression between patients and healthy controls ($P<0.05$) to develop a Combined Protein Risk Score (CPRS). The CPRS exhibited excellent diagnostic capability for detecting silicosis, with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.808 [95% confidence interval (*CI*): 0.763, 0.853; Model 1]. The traditional model including age and years of exposure, also demonstrated good diagnostic performance (ROC=0.844, 95% *CI*: 0.801, 0.888; Model 0). When CPRS and clinical variables were incorporated into Model 2, the diagnostic performance significantly improved ($P=1.546\times10^{-5}$), reaching an AUC of 0.915 (95% *CI*: 0.873, 0.937) (Figure 2A). To validate the robustness of our findings, we performed a 10-fold cross-validation with 500 repetitions. The average AUC was 0.846 (95% *CI*: 0.802, 0.890) for Model 0, 0.809 (95% *CI*: 0.764, 0.854) for Model 1, and 0.908 (95% *CI*: 0.877, 0.939) for Model 2 (Figure 2B).

DISCUSSION

Our study identified 16 novel loci and eight plasma proteins associated with susceptibility to silicosis, among which two independent SNPs (rs6677666 and rs2272528) demonstrated particularly significant biological relevance. The rs6677666 variant, located within the intron of the WLS gene, may influence pulmonary fibrosis through Wnt/β-catenin signaling activation and epithelial–mesenchymal transition (EMT). Similarly, rs2272528 in COL4A4 promotes fibrosis through collagen IV overexpression and extracellular matrix (ECM) accumulation (*5*). The ECM provides structural support as an extracellular scaffold, whereas EMT involves transformation of epithelial cells into mesenchymal cells. Both of these processes contribute to tissue fibrosis when dysregulated. Although the specific functions of RNU4-43P and LINC03000 in silicosis are uncharacterized, emerging evidence indicates that

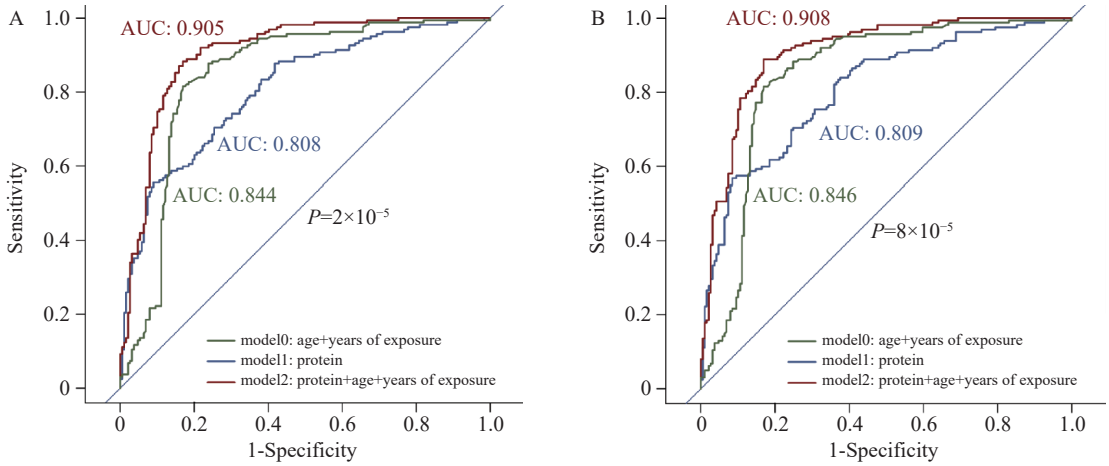


FIGURE 2. ROC curves of the protein diagnostic model. (A) ROC curves of the protein diagnostic model; (B) ROC curves after the robustness test.
Abbreviation: ROC=receiver operating characteristic; AUC=area under the curve.

pseudogenes can act as competitive endogenous RNAs to regulate parental genes (*6*), whereas long non-coding RNAs (lncRNAs) can directly modulate profibrotic pathways such as TGF-β (*7*). Thus, we speculated that RNU4-43P and LINC03000 may influence silicosis susceptibility by similarly affecting mRNA splicing stability or pro-fibrotic gene networks. Using MR analysis with pQTLs as instrumental variables, we established causal relationships between eight proteins and silicosis. Subsequent validation identified five significantly differentially expressed proteins (MMP12, EGF, Gal_9, GZMA, and ICOSLG) between patients and controls. These biomarkers are biologically relevant: MMP12 contributes to elastin degradation and fibrosis, with elevated levels confirmed in silica-exposed workers (*8*); EGF promotes pulmonary epithelial/fibroblast proliferation (*9*); Gal_9 mediates immune regulation through Tim-3 interactions; GZMA reflects cytotoxic immune activity (*10*); and ICOSLG modulates $CD4^+$ T cell activation (*11*), with emerging roles in fibrotic disorders. Their functions align with the pathogenic mechanisms of silicosis, including macrophage activation, immune cell recruitment, and chronic inflammation.

Current silicosis surveillance methods (chest radiography, high-resolution CT (HRCT), and pulmonary function tests) primarily detect advanced disease, highlighting the critical need for early biomarkers. Our multi-omics approach integrated GWAS using the ASA chip with Olink-targeted proteomics, enabling pQTL analysis to identify genetic regulators of protein expression. This strategy identified five protein bio-CPRS markers incorporated into a composite risk prediction score for early diagnosis before radiographic changes occurred. This model offers substantial clinical utility for early intervention and personalized management, particularly in patients with comorbidities such as tuberculosis or lung cancer (*12*).

Notable strengths of this study include our novel multi-omics framework combining GWAS and CPRS proteomics, a systematic analytical pipeline (pQTL, MR), and population-specific genetic profiling. However, certain limitations exist: although patients and controls were matched for silica exposure duration, residual confounding by factors such as age or smoking cannot be ruled out. Additionally, although MR suggested causal relationships between proteins and silicosis, sample size constraints necessitated validation in larger cohorts.

In conclusion, we identified two functionally significant genetic variants and five protein biomarkers that are mechanistically linked to silicosis pathogenesis. The CPRS is a robust tool for early detection in high-risk populations. By integrating genetic and proteomic markers, this study deepens our understanding of silicosis mechanisms and offers a foundation for improved public health strategies, targeted therapies, and personalized prevention approaches for this irreversible occupational disease.

Conflicts of interest: No conflicts of interest.

Acknowledgments: The medical institutions and staff who participated in this survey.

Ethical statement: Approved by the Ethics Committee of Jiangsu Provincial Center for Disease Prevention and Control (Approval No. JSJK2022-

B002-01).

Funding: Supported by the Jiangsu Provincial Social Development Program of the Key R&D Project (BE2022803), Natural Science Foundation of Jiangsu (BK20201485), Jiangsu Provincial Key Medical Discipline (ZDXK202249), and Scientific Research Project of Jiangsu Health Commission (M2022085).

doi: 10.46234/ccdcw2025.270

[#] Corresponding authors: Zhengdong Zhang, zdzhang@njmu.edu.cn; Baoli Zhu, zhubl@jscdc.cn.

[1] Department of Genetic Toxicology, The Key Laboratory of Modern Toxicology of Ministry of Education, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China; [2] Jiangsu Provincial Center for Disease Control and Prevention, Nanjing City, Jiangsu Province, China.

Submitted: August 06, 2025
Accepted: November 24, 2025
Issued: December 19, 2025

REFERENCES

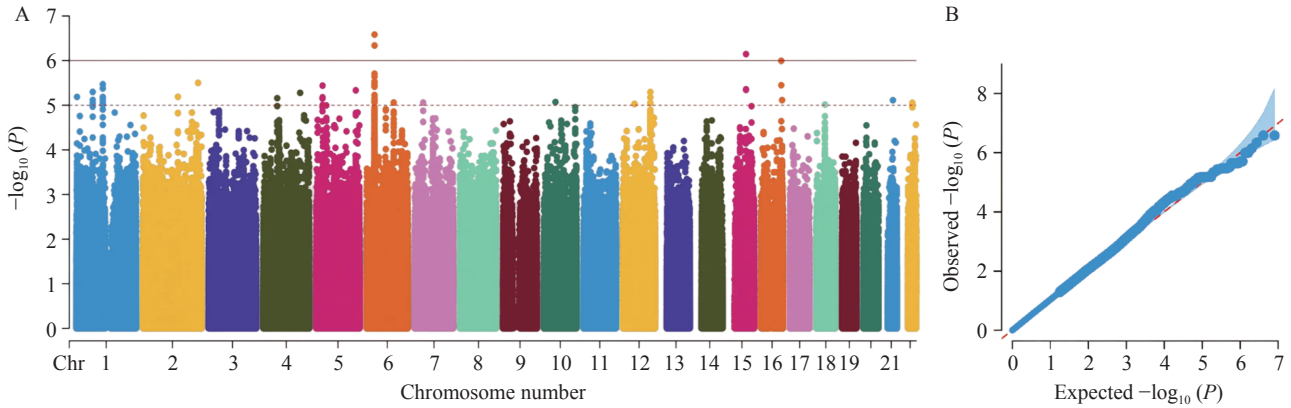
1. Li T, Yang XY, Xu H, Liu HL. Early identification, accurate diagnosis, and treatment of silicosis. Can Respir J 2022;2022(1):3769134. https://doi.org/10.1155/2022/3769134.
2. Chu MJ, Ji XM, Chen WH, Zhang RY, Sun CQ, Wang T, et al. A genome-wide association study identifies susceptibility loci of silica-related pneumoconiosis in Han Chinese. Hum Mol Genet 2014;23(23):6385 – 94. https://doi.org/10.1093/hmg/ddu333.
3. Cao ZJ, Song MY, Liu Y, Pang JL, Li ZG, Qi XM, et al. A novel pathophysiological classification of silicosis models provides some new insights into the progression of the disease. Ecotoxicol Environ Saf 2020;202:110834. https://doi.org/10.1016/j.ecoenv.2020.110834.
4. Wu YT, Xu HW, Mao LP, Zhao RR, Chu JF, Huang LL, et al. Identification of novel pQTL-SNPs associated with lung adenocarcinoma risk: a multi-stage study. Cancer Med 2024;13(17):e70247. https://doi.org/10.1002/cam4.70247.
5. Cao ZL, Xiao QL, Dai XN, Zhou ZW, Jiang R, Cheng YS, et al. circHIPK2-mediated σ-1R promotes endoplasmic reticulum stress in human pulmonary fibroblasts exposed to silica. Cell Death Dis 2017;8(12):3212. https://doi.org/10.1038/s41419-017-0017-4.
6. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A *ceRNA* hypothesis: the Rosetta Stone of a hidden RNA language? Cell 2011;146(3):353-8. http://dx.doi.org/10.1016/j.cell.2011.07.014.
7. Zhang XL, Hong RY, Chen WQ, Xu MW, Wang LF. The role of long noncoding RNA in major human disease. Bioorg Chem 2019;92:103214. https://doi.org/10.1016/j.bioorg.2019.103214.
8. Chaudhuri R, McSharry C, Brady J, Donnelly I, Grierson C, McGuinness S, et al. Sputum matrix metalloproteinase-12 in patients with chronic obstructive pulmonary disease and asthma: relationship to disease severity. J Allergy Clin Immunol 2012;129(3):655 – 63.e8. https://doi.org/10.1016/j.jaci.2011.12.996.
9. Miao RM, Ding BM, Zhang YY, Xia Q, Li Y, Zhu BL. Proteomic profiling change during the early development of silicosis disease. J Thorac Dis 2016;8(3):329 – 41. https://doi.org/10.21037/jtd.2016.02.46.
10. Ahamed MT, Forshed J, Levitsky A, Lehtiö J, Bajalan A, Pernemalm M, et al. Multiplex plasma protein assays as a diagnostic tool for lung cancer. Cancer Sci 2024;115(10):3439 – 54. https://doi.org/10.1111/cas.16300.
11. Wang SD, Zhu GF, Chapoval AI, Dong HD, Tamada K, Ni J, Chen LP. Costimulation of T cells by B7-H2, a B7-like molecule that binds ICOS. Blood 2000;96(8):2808 – 13. https://doi.org/10.1182/blood.V96.8.2808.
12. Zhang TT, Wang YY, Sun YL, Song MY, Pang JL, Wang MY, et al. Proteome, lysine acetylome, and succinylome identify posttranslational modification of STAT1 as a novel drug target in silicosis. Mol Cell Proteomics 2024;23(6):100770. https://doi.org/10.1016/j.mcpro.2024.100770.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. Baseline characteristics of participants in the case–control study of plasma proteins for silicosis.

Characteristic	Case (*n*=163)	Control (*n*=189)	*P*
Sex (*n*, %)			0.312
Female	2 (1.2)	0 (0)	
Male	161 (98.8)	189 (100.0)	
Age [mean (SD)]	72.0 [69.0, 75.0]	67.0 [64.0, 71.0]	<0.001
Smoking status (*n*, %)			0.707
Never	43 (26.4)	54 (28.6)	
Former	63 (38.7)	65 (34.4)	
Current	57 (35.0)	70 (37.0)	
Pack-year [median (IQR)]	15.0 (0.0, 25.5)	10.0 (0.0, 25.0)	0.126
Drinking status (*n*, %)			0.651
Never	40 (24.5)	39 (20.6)	
Former	47 (28.8)	60 (31.7)	
Current	76 (46.6)	90 (47.6)	
Cumulative dust exposure (mg/m^3·y)	1,688.0 (1,506.1, 1,857.2)	1,648.1 (1,253.4, 1,788.1)	0.014
Use of personal protective equipment (PPE)			0.597
No	15 (9.2)	23 (12.2)	
Yes	147 (90.2)	164 (86.8)	

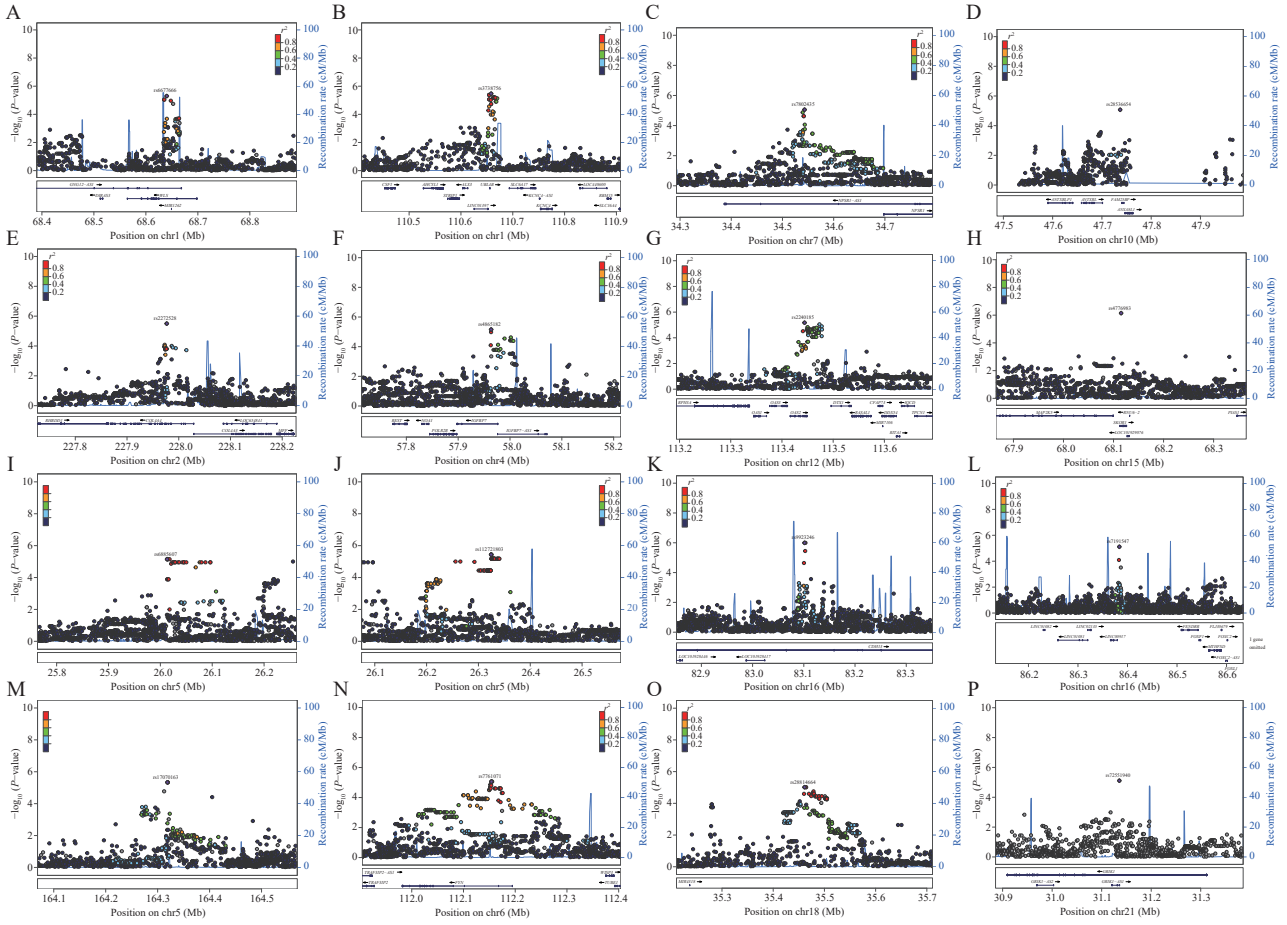
Abbreviation: SD=standard deviation; IQR=interquartile range; PPE=personal protective equipment.



SUPPLEMENTARY FIGURE S1. Results of GWAS for the silicosis. (A) Manhattan plot for the silicosis GWAS; (B) Quantile–quantile plot for the silicosis GWAS.

Note: In the Manhattan plot, different colors represent different chromosomes for easier distinction. Each point represents an SNP, i.e., a genetic variant site. A higher point position indicates a stronger statistical association between the locus and the risk of developing silicosis. (A) used to visually display which chromosomal regions contain genetic loci significantly associated with silicosis. A smaller P corresponds to a larger -$\log_{10}$(P), indicating a stronger association between the locus and silicosis. (B) used to assess the quality of the GWAS analysis and to check for systematic bias or true genetic effects. The red dashed line represents the expected distribution line. If no true associations exist, the observed values should largely coincide with the expected values. The blue line represents the observed distribution line. If the blue line deviates significantly from the red dashed line in the region of low P (upper right corner), it indicates the presence of true genetic signals (i.e., SNPs associated with the disease). If the blue line largely overlaps with the red dashed line overall, it suggests that the analysis may not have found significant associations. If the blue line shows excessive deviation in the low P region, it might suggest the presence of confounding factors such as population stratification.

Abbreviation: GWAS=genome-wide association study; SNP=single-nucleotide polymorphism.



SUPPLEMENTARY FIGURE S2. Regional plots showing the chromosomal positions of 16 independent risk loci identified in the silicosis GWAS. (A) Position on chr1 (Mb); (B) Position on chr1 (Mb); (C) Position on chr7 (Mb); (D) Position on chr10 (Mb); (E) Position on chr2 (Mb); (F) Position on chr4 (Mb); (G) Position on chr12 (Mb); (H) Position on chr15 (Mb); (I) Position on chr5 (Mb); (J) Position on chr5 (Mb); (K) Position on chr16 (Mb); (L) Position on chr16 (Mb); (M) Position on chr5 (Mb); (N) Position on chr6 (Mb); (O) Position on chr18 (Mb); (P) Position on chr21 (Mb).
Abbreviation: GWAS=genome-wide association study.

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).

CCDC weekly

Vol. 7 No. 51 Dec. 19, 2025

Responsible Authority
National Disease Control and Prevention Administration

Sponsor
Chinese Center for Disease Control and Prevention

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

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