Expanding mountain-type zoonotic visceral leishmaniasis in China: a multi-source data-driven investigation

Zhengbin Zhou,[a,h] Yanfeng Gong,[b,h] Lulu Huang,[a] Zhongqiu Li,[a] Yuwan Hao,[a] Yuanyuan Li,[a] Haobo Ni,[a] Limin Yang,[a] Xinyi Wang,[a] Canjun Zheng,[c] Ying Liu,[d] Xiaodong Tian,[e] Wenting Wu,[f] Qin Liu,[a] Yi Zhang,[a] Shang Xia,[a] Junhu Chen,[a] Yibiao Zhou,[b] Robert Berquist,[g] Xiaonong Zhou,[a] Shizhu Li,[a,*] and Qun Li[c,**]

[a]National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Chinese Center for Tropical Diseases Research, National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, NHC Key Laboratory of Parasite and Vector Biology, WHO Collaborating Centre for Tropical Diseases, National Center for International Research on Tropical Diseases, Ministry of Science and Technology, Shanghai 200025, China
[b]Department of Epidemiology, School of Public Health, Key Laboratory of Public Health Safety (Fudan University), Ministry of Education, Fudan University, Shanghai, China
[c]Chinese Center for Disease Control and Prevention, National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, Beijing 102299, China
[d]Department of Parasite Disease Control and Prevention, Henan Center for Disease Control and Prevention, Zhengzhou, Henan, China
[e]Shanxi Provincial Center for Disease Control and Prevention, Taiyuan, Shanxi, China
[f]Institute for Infectious Disease and Endemic Disease Control, Beijing Center for Disease Prevention and Control, Beijing, China
[g]Ingerod, Brastad, Sweden (formerly at the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR)), World Health Organization, Geneva, Switzerland

The Lancet Regional Health - Western Pacific 2026;70: 101858

Published Online xxx
https://doi.org/10.1016/j.lanwpc.2026.101858

Summary

Background Mountain-type zoonotic visceral leishmaniasis (MT-ZVL) is undergoing rapid re-emergence in central China and demonstrates significant potential for further dissemination. The aim of this study was to provide information on optimal control strategies by estimating the disease's current distribution and potential for spread into new areas.

Methods We used a three-pronged approach: (1) active field surveillance in 57 potential-risk counties in 2022 to investigate the distribution of the *Phlebotomus chinensis* (*P. chinensis*) vector and canine infections based on serological positivity; (2) study of annual surveillance reports of *P. chinensis* distribution and canine infection in 42 counties in Henan, Shanxi and Beijing (2023–2024); and (3) passive surveillance via China's National Notifiable Disease Surveillance System (NNDSS) revealing 1137 human MT-ZVL cases for the 2019–2024 period. We constructed a database based on an innovative active/passive surveillance approach and applied the optimized ecological niche model based on incorporating environmental, biological (the *P. chinensis* vector and infectious canines) and socioeconomic covariates to predict the relative risk of transmission at the 1 × 1 km resolution across central China. Model performance was validated using 239 unique, geolocated human case occurrences (2022–2024). This allowed us to map the MT-ZVL risk and quantify the populations at-risk within stratified risk tiers.

Findings The field surveillance across 20 of 57 potential MT-ZVL risk counties captured 7552 *P. chinensis* specimens (mean density: 13.7 sandflies/trap/night), with a high peak in Jiexiu County, Shanxi Province (499.8). In 14 counties with confirmed *P. chinensis* presence, canine infections were recorded in 9 counties giving a mean seroprevalence of 2.7% (44/1648), with the highest proportion at 16.2% (21/130) seen in Yuzhou County, Henan Province. Annual surveillance reports confirmed *P. chinensis* presence in 24 of 42 monitored counties, with seropositive canines in two counties. The final training dataset included the recorded findings of 270 *P. chinensis* sandflies, 224 infected canines and 213 human case records. Extreme gradient boosting demonstrated superior predictive performance for human MT-ZVL infection risk, significantly outperforming benchmark models, with the area under the curve (AUC) of 0.950 (95% CI: 0.921–0.979) in internal validation and 0.913 (95% CI: 0.886–0.940) in external validation. Model performance was further validated by 88.3% of indigenous cases occurring within predicted high-risk zones. These high-risk zones were concentrated in hilly terrain along the Taihang, Lvliang and Qinling Mountains—primarily in central-southern Shanxi, Hebei-Shanxi border regions, eastern Shaanxi and

*Corresponding author. No. 207, Ruijin Er Road, Huangpu District, Shanghai, China.
**Corresponding author. No. 155, Changbai Road, Changping District, Beijing, China.
E-mail addresses: lisz@chinacdc.cn (S. Li), liqun@chinacdc.cn (Q. Li).
[h]These authors contributed equally to this work.

north-western Henan. This study identified 117 high-risk or very high-risk counties with a combined population of 19.24 million. Medium-risk areas were found to encompass surrounding regions across Shanxi, Shaanxi, Henan, eastern Gansu, central-western Hebei and western/northern Beijing (87 counties; 12.07 million residents), while low-risk zones covered 87 counties in eastern Gansu, northern Sichuan, north-western Henan and central Hebei mountains (15.35 million residents). The optimized multi-source data-driven model exhibited robust predictive performance, identifying 291 at-risk counties threatening about 46.67 million residents.

Interpretation This study demonstrated MT-ZVL to be a significant public health threat with rapid expansion across central China. Key findings revealed unabated endemic transmission, with the presence of *P. chinensis* and infectious canines in non-endemic areas signifying new emerging foci. By the identification of specific regions at risk for MT-ZVL, an evidence base warranting heightened public health vigilance against potential transmission has been established. Surveillance for MT-ZVL in central China and targeted interventions are strongly recommended.

Funding This work was supported by the National Natural Science Foundation of China (No. 32161143036, No. 82473688), National Key Research and Development Program (2021YFC2300800, 2024YFC2310902), the Three-Year Initiative Plan for Strengthening Public Health System Construction in Shanghai (2023–2025) (No. GWVI-11.1-12).

Keywords: Visceral leishmaniasis; Resurgence; Vulnerable population; *Phlebotomus chinensis*; Risk prediction; China

Research in context

Evidence before this study

We searched PubMed without language restrictions from database inception up to November 30th, 2025, for studies published with the terms ("Visceral leishmaniasis*" [All Fields] OR "Kala-azar" [All Fields]) AND ("geograph*" [All Fields] OR "spati*" [All Fields]). Our search yielded more than 906 studies; We found evidence of zoonotic visceral leishmaniasis re-emergence or reoccurrence in Asia, Africa, South America and Europe in naive populations after several years of low incidence or no infections; More than 50 of these studies of the risk mapping were related to the association between environmental or biological factors and mountain-type zoonotic visceral leishmaniasis (MT-ZVL) in some local endemic regions. However, no integrated model simultaneously incorporates MT-ZVL incidence, surveillance-derived biological covariates and environmental/socio-demographic variables across endemic/potential risk regions to simulate county-level MT-ZVL risk heterogeneity and predict transmission receptivity beyond known endemic zones on a larger spatial scale.

Added value of this study

To our knowledge, this represents the first integrated modeling framework—combining multi-source data-driven risk forecasting with vulnerable population assessment—that incorporates MT-ZVL infection data, biologically relevant covariates from active surveillance, and environmental/socio-demographic variables. Our model simulates county-level heterogeneity in MT-ZVL risk across all at-risk areas of China while predicting transmission receptivity beyond currently identified risk zones. This study demonstrates MT-ZVL as a significant public health threat exhibiting rapid expansion throughout central China. Key findings reveal unabated endemic transmission, with *Phlebotomus chinensis* and infectious canines detected in non-endemic areas indicating emerging foci. The presence of *P. chinensis* in high-latitude regions likely drives northward expansion of MT-ZVL transmission. The optimized multi-source model demonstrated robust predictive performance, identifying 291 at-risk counties threatening approximately 46.67 million residents in China.

Implications of all the available evidence

Our findings suggest that MT-ZVL epidemics will continue to occur due to the geographical heterogeneity associated with MT-ZVL spread or recurrence in China. Our model can be used to identify specific regions at risk for MT-ZVL. These results will be useful in informing public health interventions to anticipate and prevent future MT-ZVL in China.

Introduction

Visceral leishmaniasis (VL), also known as kala-azar, is a life-threatening zoonotic parasitic disease caused by *Leishmania* protozoa.[1,2] The disease is transmitted through bites of infected female sandflies (*Phlebotomus*), with parasite reservoirs maintained in canids, rodents and humans.[1] Untreated human cases exhibit a case fatality rate exceeding 95% within 1–2 years after the onset of symptoms.[3] Globally, VL represents the second most lethal parasitic disease after malaria–endemic in

88 countries, it imposes a substantial global health burden, with 50,000–90,000 new cases reported annually.[1,3] Recognized by the World Health Organization (WHO) as a neglected tropical disease (NTD), VL is targeted for global elimination as a public health problem under WHO's NTD Roadmap 2021–2030,[4] particularly affecting impoverished regions.

In China, VL is epidemiologically classified into three forms: anthroponotic visceral leishmaniasis (AVL), mountain-type zoonotic visceral leishmaniasis (MT-ZVL) and desert-type zoonotic visceral leishmaniasis (DT-ZVL).[2,5] *P. chinensis* (a semi-wild subspecies) is the sole confirmed vector of MT-ZVL. Domestic dogs serve as the primary reservoir of the parasite (*Leishmania infantum*) that exhibits high infection rates, with humans acting as accidental hosts.[2,5] MT-ZVL resurgence can occur when *L. infantum* is introduced by infectious canines into receptive, densely populated areas characterized by abundant *P. chinensis* populations and insufficient prevention measures.[6,7] Although MT-ZVL was eliminated from most parts of China by the 1980s, the past two decades have witnessed a swift resurgence and geographical spread of MT-ZVL in the western and central mountainous regions in China[7–10] highlighting the urgent need to better understand both current MT-ZVL distributions and their potential to spread into new areas. Significant MT-ZVL resurgence, documented over the last five years in Shanxi, Shaanxi, Hebei, Henan and Beijing provinces together with the established distribution of *P. chinensis* in the hilly regions[10] indicates a substantial potential for further dissemination of MT-ZVL within China.

Expansion of the disease would increase the financial burden on health systems and exacerbate the detrimental impact on population health.[1,10] However, updating current control strategies for MT-ZVL requires an improved evidence base for identifying the most vulnerable populations. Although risk mapping for MT-ZVL occurrence in areas at risk of transmission has been done at provincial and sub-provincial levels previously,[8,11] the spatial distribution of the disease within these regions remains poorly defined since available *P. chinensis* distribution data for areas outside defined risk zones are mostly decades-old.[7,8,11] While *P. chinensis* presence constitutes a necessary condition for transmission, numerous other factors contribute to the establishment and maintenance of transmission cycles.[2] Notably, the emergence of indigenous human cases in potential risk areas signifies long-established, persistent endemic transmission cycles of the parasite characterized by extensive *P. chinensis* colonization and high infection prevalence within canine reservoirs.[8,9,12] Therefore, the broad-scale implementation of generalized control campaigns across extensive geographical regions confronts formidable operational and economic barriers related to resource allocations, which impose substantial and sustained socioeconomic burdens on affected communities.[13,14]

Pre-emptive deployment of precision interventions within epidemiologically demarcated transmission hotspots prior to disease resurgence represents an essential strategic paradigm for disrupting comparatively unstable transmission chains in their incipient stages. To address this imperative, we conducted a large-scale active surveillance of *P. chinensis* distribution and canine seroprevalence across potential risk zones—the first comprehensive epidemiological study of its kind in China. This initiative was based on the study of critical biological covariates essential for developing a refined, high-sensitivity predictive risk model to guide intervention strategies. Unlike previous studies relying solely on passive case surveillance with inherent reporting delays,[7,8,11] our integrated surveillance data combined passive (case reporting) and active surveillance (systematic field monitoring of vector distribution and reservoir hosts with pathogen detection), as this approach would enable proactive risk identification prior to human MT-ZVL case emergence,[11] thereby enhancing data completeness and timeliness for improved transmission risk assessment of MT-ZVL. We employed an integrated modelling framework—multi-source data-driven risk forecasting and vulnerable population assessment—incorporating MT-ZVL infection data; biologically relevant covariates from active surveillance; and environmental and socio-demographic variables, a model not only simulating county-level heterogeneity in MT-ZVL risk across all at-risk areas of China but also predicting receptivity to MT-ZVL transmission in regions beyond currently identified risk zones.

Methods

Study area

Based on previous distributions of *P. chinensis* and recently reported VL cases, 10 provincial-level administrative divisions (PLADs)—Gansu, Sichuan, Shanxi, Shaanxi, Henan, Hebei, Qinghai, Ningxia, Liaoning and Beijing—were selected as study areas.[2] These areas span central China (26°N–44°N, 89°E–126°E) encompass 1002 counties and feature a west-high to east-low gradient of the topography, with elevations varying from –4 to 5068 m. The mean annual temperature ranges from –2 °C to 18 °C and the precipitation decreases from southeast to northwest (200–1800 mm, respectively).

Data sources

Field survey data (active surveillance)

Active field surveillance was prioritized in six of the 10 PLADs—Shanxi, Shaanxi, Henan, Hebei, Gansu, and Sichuan—based on their historical role as the most severely affected regions, accounting for all indigenous MT-ZVL cases reported between 2005 and 2020. To guide field surveys across 812 counties in six endemic

provinces, we developed a three-stage sampling framework. First, using K-means clustering of ecological variables (e.g., elevation, NDVI, temperature), we stratified the region into high-, medium-, and low-risk clusters. Second, we applied screening criteria—excluding areas outside the sandfly's elevational niche (50–2000 m), highly urbanized zones, and endemic counties (2015–2020)—yielding a sampling frame of 629 "potential risk" counties. Finally, we implemented stratified random sampling with higher intensity in higher-risk strata, selecting 57 counties for entomological and canine surveys during peak sandfly activity periods (June–July 2022). Further details on the construction of the sampling framework are provided in the Appendix.

The protocol involved three livestock-raising villages per county, with ≥2 light traps per village for continuous 3-day collections of *P. chinensis*. In a subset of villages with confirmed *P. chinensis* presence, canine seroprevalence surveys were conducted, collecting ≥80 canine serum samples per county. Serological testing for *L. infantum* antibodies was performed using the rk39 RDT (Kalazar Detect™). In 2023–2024, surveillance activities guided by the China CDC extended this approach to 23 counties/districts in Henan Province, 14 districts in Beijing, and 5 counties in Shanxi Province.

Web-based notifiable disease data (passive surveillance)

All epidemiological data related to MT-ZVL cases were obtained from the National Notifiable Diseases Reporting System (NNDSS) of China CDC. Townships where indigenous MT-ZVL cases had been reported during the 2019–2024 period were defined as occurrence sites, with longitude and latitude coordinates measured for each site. From this collected information, suspected and duplicate cases were excluded. Descriptive epidemiological analysis of the three-dimensional distribution of VL was conducted using Microsoft Excel 2016 software (Microsoft Corp., Redmond, WA, USA).

Modeling explanatory covariates

This study employed a stacked modeling strategy, in which predictions from the underlying models—specifically, the vector suitability model and the canine infection risk model—were used as covariates in the top-level model for predicting human infection risk. This multi-layer approach enables the integration of diverse sources of biological information. A range of environmental and biological factors potentially influencing MT-ZVL transmission were compiled as 1 × 1 km gridded surfaces and considered as candidate explanatory covariates (Table 1). Predictive distributions of *P. chinensis* and Leishmania-infected canines were generated using multiple machine learning models and subsequently incorporated as biological covariates. Detailed procedures for constructing these biological factors are provided in the Appendix.

Modeling approach

We utilized a machine learning model that has previously demonstrated efficacy in risk mapping of vector-borne diseases.[15] The MT-ZVL infection risk model was developed (model training and testing) using human case data from 2019 to 2021 and subsequently validated with case records from 2022 to 2024. All modeling data—for *P. chinensis* distribution, infected canine, and human MT-ZVL cases—were derived from the seven study provinces (Gansu, Sichuan, Shaanxi, Shanxi, Henan, Hebei, and Beijing). A comprehensive summary of the data sources, time periods, and

Factor	Parameter	Abbreviation	Time	Spatial resolution	Data source
Ecology	Normalized difference vegetation index	NDVI	2019–2024	1 km	MODIS Satellite data[c]
	Distance to waterway	DS_WW (km)	2019–2024	100 m	WorldPop[d]
	Distance to nature reserves and wilderness	DS_NW (km)	2019–2024	100 m	
Topography	Elevation	Elv (m)	2007	100 m	
	Slope	Slp (°)	2007	100 m	
Climate	Annual average temperature	T_{ave} (°C)	2019–2024	1 km	NESSDC
	Annual minimum temperature[a]	T_{min} (°C)	2019–2024	1 km	
	Annual maximum temperature[b]	T_{max} (°C)	2019–2024	1 km	
	Annual total precipitation	P_{tot} (mm)	2019–2024	1 km	
	Relative humidity	Hum (1%)	2019–2024	1 km	
Socioeconomic	Night-time light index	NTLI	2019–2024	1 km	WorldPop[d]
Vector suitability	Predicted distribution of *Phlebotomus chinensis*	*P. chinensis*		1 km	This study
The reservoirs	Predicted distribution of canine infected with *Leishmania*	Can.		1 km	This study

[a]In January. [b]In July; NESSDC = National Earth System Science Data Center. [c]MODIS (https://modis.gsfc.nasa.gov/). [d]WorldPop (https://hub.worldpop.org/).

Table 1: **Environmental variables affecting the distribution of mountain-type zoonotic visceral leishmaniasis.**

temporal representativeness of the predictions for each model is available in Appendix Supplementary Tables S1 and S3. The overall MT-ZVL occurrence model equation was constructed as follows:

$$OCC_{MT-ZVL,i} \sim Bernoulli\left(P\left(OCC_{MT-ZVL,i}\right)\right) \quad \text{(Eq. 1)}$$

$$logit\left(P\left(OCC_{MT-ZVL,i}\right)\right) \sim f\left(Eco_i, Top_i, Cli_i, Soc_i, Phleb_i, Can_i\right) \quad \text{(Eq. 2)}$$

where $P(OCC_{MT-ZVL,i})$ is the location-specific probability that a record i is an occurrence record of an indigenous MT-ZVL case, while f denotes the nonlinear function of all the covariates $\left(Eco_i = NDVI, Top_i = topographic\ variable, Cli_i = climate\ variable, Soc_i = NTLI, Phleb_i = P.\ chinensis, Can_i = Infected\ canine\right)$. The climate variables comprised annual average temperature (Tave), annual minimum temperature (Tmin), annual maximum temperature (Tmax), annual total precipitation (Ptot), and relative humidity (Hum). We characterized the machine learning function using six algorithms: logistic regression (LR),[16] support vector machine (SVM),[17] random forest (RF),[18] multilayer perceptron machine (MLP),[19] light gradient boosting machine (LGB),[20] and extreme gradient boosting (XGB).[21] Further details on the algorithm of the models are provided in the Appendix.

Model validation

To ensure the independence of training and test data, we employed 10-fold cross-validation to evaluate model performance.[22,23] In each iteration, one fold (10%) of training data were for validation, the remaining 90% for model calibration. By cycling through the dataset, each fold segment was utilized for validation for a dataset only once. Finally, the model performance metrics were aggregated leveraging the area under the curve (AUC) of the receiver-operating characteristic (ROC) as a measure of the model's goodness-of-fit (AUC quantitatively evaluates the model's proficiency in distinguishing between presence and background points[24]). The optimal probability threshold for classifying predicted risks was determined as 0.403 based on the Youden index, and this threshold was used to calculate sensitivity and specificity. For external validation, we used 239 indigenous MT-ZVL cases reported between 2022 and 2024, which were independent of the 2019–2021 training data. We computed AUC, sensitivity and specificity to gauge the model performance on this unseen data. In addition to discrimination, we assessed the calibration of the XGBoost model using the Brier score and calibration plots on the external validation dataset.

Vulnerable population assessment

To ensure prediction realism and prevent overestimation, we implemented post-hoc spatial masking to nullify risk in areas with biologically implausible transmission potential,[25] particularly urban cores (e.g., Lanzhou, Chengdu) where habitat degradation eliminated suitable *P. chinensis* environments. We established a unified 0–1 risk scale divided into five equal-interval tiers: 0.00–0.20 (very low), 0.21–0.40 (low), 0.41–0.60 (moderate), 0.61–0.80 (high), and 0.81–1.00 (very high). To ensure the rationality of this classification, we counted the number and proportion of external validation cases falling into each grade.

County risk levels were classified as: low-risk, moderate-risk, high-risk, and very high-risk based on the maximum risk value within each county's boundaries. The maximum value principle mainly takes into account the risk transmission: the existence of any maximum-risk area within a county indicates that the transmission chain formed by the vector and the canine host may be established in the county; even if the maximum-risk area is small, it may become foci for the spread of the epidemic. Rural populations within risk county were used to quantify exposed individuals, as MT-ZVL transmission is ecologically constrained to: (1) rural areas with altitudes of 400–1600 m where sandfly vectors thrive, and (2) agricultural zones maintaining canine reservoir hosts. Rural population data were from China's Seventh National Census (https://www.stats.gov.cn/sj/pcsj/rkpc/d7c/).

All statistical analyses and visualizations were performed using R 4.3.2 (R Core Team, Vienna, Austria) and ArcGIS v. 10.1 (ESRI Inc., Redlands, CA, USA).

Ethics

Ethics approval was granted by the Ethics Committee of the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention; Chinese Center for Tropical Diseases Research (approval No. IPD-2022-039, 10 August 2022). The study involved blood sampling and antibody testing of dogs. Informed consent was obtained from all dog owners prior to sample collection.

Role of the funding source

The funder of the study had no role in study design, data analysis, data interpretation, or writing of this paper.

Results

Field surveys of *P. chinensis* and canines

In the 57 sampled counties, surveillance confirmed presence of *P. chinensis* in 20 of them, where 7552 specimens were captured at a mean density of 13.7 per trap/night. As seen in Fig. 1, Jiexiu City, Shanxi, had the highest density (499.8 *P. chinensis*/trap/night). The infested counties were found to be clustered in the mountainous Loess Plateau regions of Shanxi, along the Shanxi–Henan border and the Gansu–Shaanxi

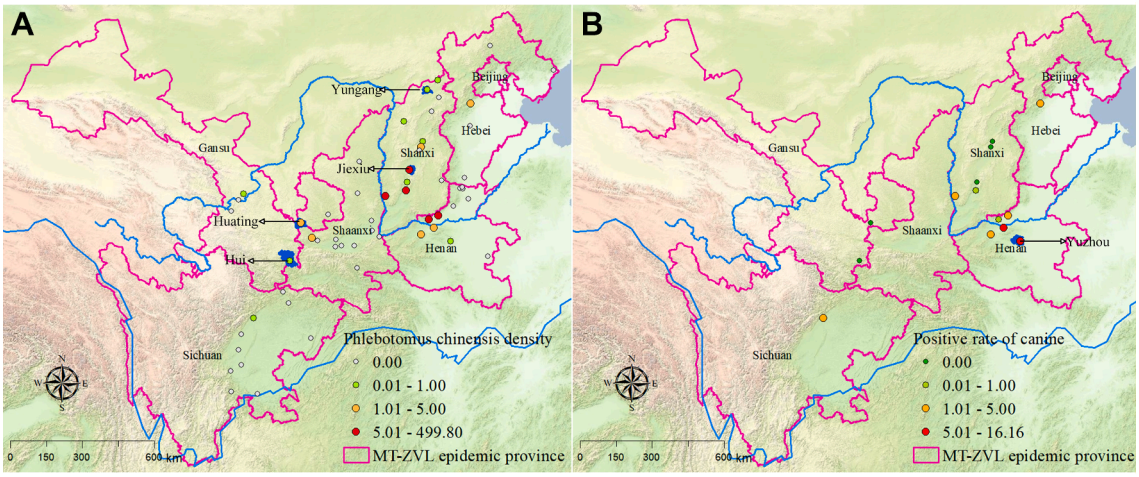

***Fig. 1:* Summary of the results of *P. chinensis* surveys using light traps (A) and rk39 RDT results for infected canines (B).** Key sites are annotated: Jiexiu City (Shanxi), where the highest sandfly density was recorded (499.8 per trap-night); Yuzhou City (Henan), with the highest canine seroprevalence (16.2%); and three newly identified *P. chinensis* distribution areas—Huating City, Hui County, and Yungang District.

border. Among these 20 counties, 17 had historical records of this vector species and three were designated newly identified endemic areas. These novel detection sites clustered predominantly within the high-altitude zones of eastern Gansu Province and the high-latitude mountainous territories of northern Shanxi Province indicating potential expansion of the *P. chinensis* distribution. Canine serological surveillance conducted on 1648 dogs in 14 counties with confirmed presence of *P. chinensis* (Shanxi: five counties; Henan: five counties; Gansu: two counties; Hebei: one county; Sichuan: one county). Of these, 44 dogs were found to be seropositive with a mean seroprevalence of 2.7%. These seropositive canines were distributed across 9 counties: Henan (five counties), Shanxi (two), Hebei (one), and Sichuan (one), with Yuzhou City, Henan showing the highest seroprevalence at 16.2% (21/130). The remaining 6 counties (Shanxi: 3; Gansu: 2; Shaanxi: 1) had dog populations but could not be surveyed due to operational constraints at the local level, including personnel shortages, competing priorities, and logistical difficulties. Provincial CDC surveillance demonstrated a distinct restriction of *P. chinensis* by elevation. In Henan, *P. chinensis* was detected in all 14 north-western mountainous/hilly counties but absent in 9 counties in other regions. Likewise in Beijing, *P. chinensis* was present exclusively in all 5 hilly districts, while absent from 9 urban districts. Additional data of the year 2023 emanating from Lvliang in Shanxi confirmed the distribution of *P. chinensis* across 5 counties/districts, covering 7 townships (Fig. 1 and Supplementary Table S1).

Epidemiological characteristics of the human cases

Between 2019 and 2024, China reported a total number of 1137 cases of MT-ZVL showing a consistently increasing trend, with case numbers moving from 121 cases in 2019 to 237 in 2024. This represents a 95.9% cumulative increase over the surveillance period, with 2024 recording the highest annual incidence to date. MT-ZVL also showed pronounced geographic heterogeneity, with 90.3% of cases (n = 1027) concentrated in provincial-level administrative divisions (PLADs) of four endemic provinces: Shanxi (n = 595, 52.3%), Henan (n = 152, 13.4%), Shaanxi (n = 140, 12.3%) and Gansu (n = 140, 12.3%). In total, 124 endemic counties were identified in seven PLADs and 46 of them (37.1%) were identified as counties in which MT-ZVL had re-emerged between 2019 and 2024, with 111 indigenous cases (Supplementary Table S7. Glossary of key terms) reported. This pattern illustrates the northward spatio-temporal expansion of MT-ZVL: following more than three decades of case-free status, indigenous cases progressively re-emerged in Henan (2016), Hebei (2019), and Beijing (2022). Consequently, the latitudinal distribution of cases shifted markedly northward, from 35.9°N in 2016 to 40.5°N in 2022—a geographic expansion of approximately 500 km. Furthermore, these cases showed distinct demographic patterns, with agricultural workers representing the majority (n = 588, 51.7%) and children comprising a significant vulnerable subgroup (n = 243, 21.4%) (Supplementary Table S2).

Predictability of MT-ZVL human infection risk

XGB outperformed the other algorithms used for model development achieving AUCs of 0.954 (95% CI: 0.929–0.979) in training, 0.950 (95% CI: 0.921–0.979)in testing, and 0.913 (95% CI: 0.886–0.940) in external validation (Supplementary Table S2). The model demonstrated good calibration (Brier score = 0.08; Supplementary Fig. S3), supporting the reliability of the

predicted risk probabilities. The risk stratification showed strong concordance with the actual case distribution in the external validation (Fig. 2). Specifically, 88.3% of indigenous human MT-ZVL cases reported between 2022 and 2024 (the external validation) occurred within areas classified as either very high-risk (74.1%) or high-risk (14.2%) zones. This high case capture rate demonstrated that XGB-based risk stratification effectively identified MT-ZVL transmission hotspots.

To gain deeper insight into the model's predictions, we performed a SHAP (SHapley Additive exPlanations) analysis to quantify the contribution of each predictor variable. The results revealed that predicted canine infection risk and predicted *P. chinensis* suitability were the two most important contributions of MT-ZVL human infection risk, together accounting for the majority (60.07%) of the model's predictive power. Detailed results of the SHAP analysis are provided in the Appendix.

The MT-ZVL risk zones primarily radiated outward from four major mountain ranges: the Taihang, Lvliang, Qinling and Yanshan Mountains (Fig. 3). The MT-ZVL high-risk areas were primarily concentrated in central-southern Shanxi, the border areas of Hebei and Shanxi, eastern Shaanxi and north-western Henan. Medium-risk areas surrounded the high-risk zones, predominantly covering Shanxi, Shaanxi, Henan, eastern Gansu, central-western Hebei and Beijing (western and northern areas). The low-risk areas were mainly distributed in the hilly regions of eastern Gansu, northern Sichuan, north-western Henan and the central Hebei mountains.

We estimated that 46.67 million people (3.4% of China's 2020 population) reside in MT-ZVL risk areas distributed across rural regions in 291 counties (117 currently confirmed endemic counties and 174 newly identified at-risk counties) within 10 provinces (Table 2). The counties were classified into four risk levels: very high-risk and high risk (mostly areas with established transmission), medium risk (primarily regions with resurgence potential) and low risk (counties with low resurgence risk). Across these zones, 19.24 million people inhabited zones with very high-risk and high risk spanning 117 counties, with concentrations in Shanxi (41 counties), Hebei (21 counties), Henan (20 counties), Gansu (18 counties), Shaanxi (9 counties), Beijing (5 counties) and Sichuan (3 counties). Simultaneously, 12.07 million resided in medium-risk areas across 87 counties and 15.35 million in low-risk areas

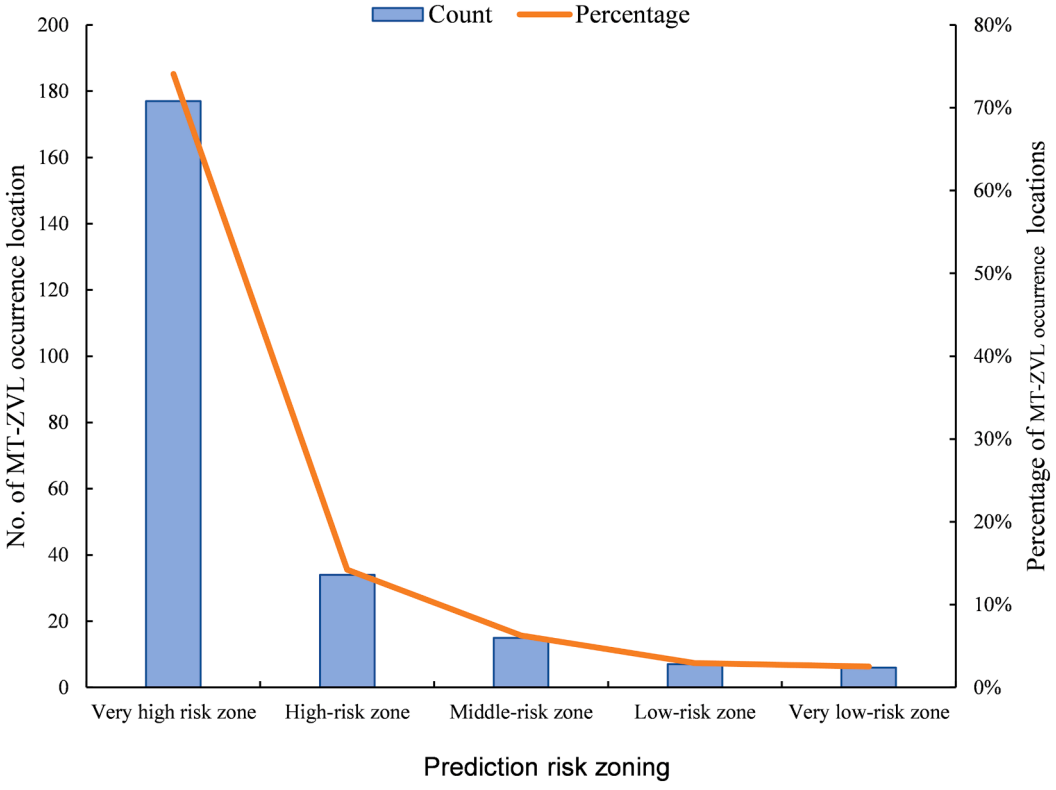

Fig. 2: Geographical concordance between XGBoost-predicted risk areas and externally indigenous case locations 2022–2024.

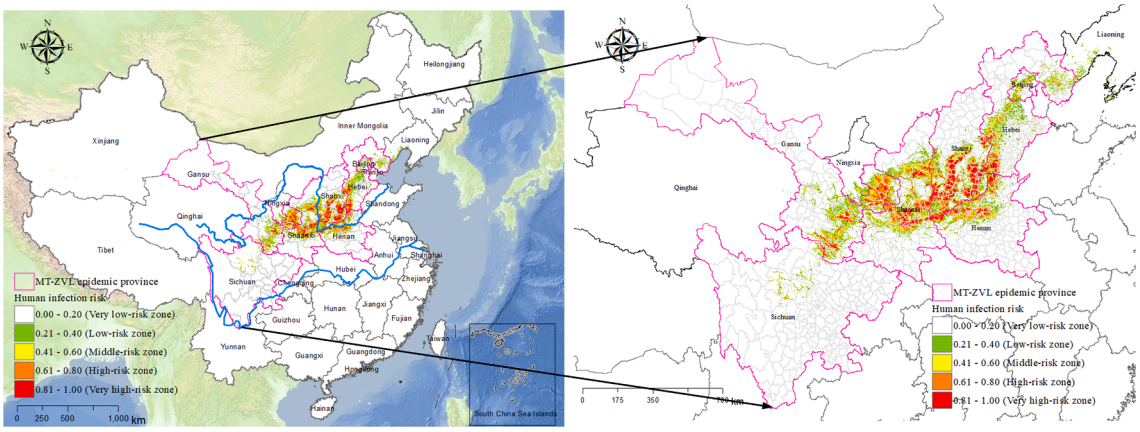

Fig. 3: XGB-based stratification of MT-ZVL human infection risk informed by canine infection risk, *P. chinensis* habitat, and environmental condition.

across 87 counties, all areas where the risk is deemed to be increasing. The province-level population estimates are given in Supplementary Table S6.

Discussion

To the best of our knowledge, this study provides the first example of an ecological niche model for forecasting the risk of MT-ZVL driven by multi-sourced data integrating active and passive surveillance data with environmental, biological and socioeconomic covariates. Active surveillance identified the 20 counties with confirmed *P. chinensis* presence, including three newly identified distribution areas, primarily in high-latitude mountainous regions, suggesting a broader distribution of this vector than previously recognized. Indigenous cases progressively re-emerged in Henan (2016), Hebei (2019), and Beijing (2022), driving a marked northward shift in their latitudinal distribution from 35.9°N to 40.5°N that underscores the northward spread of this disease.[26] This northward range expansion—a pattern consistent with global climate warming trends—highlights the potential for further spread of the disease under predicted climate scenarios.[6,27] Consistent with previous findings,[28] exceptionally high *P. chinensis* densities were recorded on the Loess Plateau indicating the impact of ecological conditions highly conducive to its proliferation. The results in Henan and Beijing confirmed the consistent colonization of *P. chinensis* on the mountainous Loess Plateau and in the peripheral highlands, whereas the adjacent plain regions showed no evidence of colonization.

Risk level	County at risk (no.)	Rural population at risk[a] (thousands)
Very high-risk zone	55	9049.88
High-risk zone	62	10191.91
Medium-risk zone	87	12070.83
Low-risk zone	87	15354.23
Total	291	46666.84

[a]Based on the data from the Seventh National Census in 2020.

Table 2: **Estimated population and number of counties at risk of MT-ZVL in central China.**

Our field surveys using rk39 RDT confirmed canine *Leishmania* infections in eight counties among the 14 counties in which *P. chinensis* was detected. However, the limited sensitivity of rk39 RDT for asymptomatic infections may have caused the true prevalence to have been underestimated.[1,24,29] Therefore, we conservatively estimate that more than 125 risk counties [$629 \times (20/57) \times (8/14)$] across the aforementioned six provinces contain infected canines. These findings suggest that multiple counties are currently in the early stages of MT-ZVL resurgence. Other historically endemic provinces—including Beijing, Liaoning, Qinghai, and Ningxia—also face analogous re-emergence threats. The passive surveillance by NNDSS revealed that MT-ZVL re-emergence has occurred in 46 historically endemic counties, and collectively reported 111 indigenous cases between 2022 and 2024. This resurgence probably relates to ecological shifts in natural foci. For instance, rural depopulation has created abandoned fields and dwellings, which are optimal habitats for *P. chinensis* and its reservoir hosts, and canine infections in guard, herding and pet dogs allow its zoonotic transmission.[10] This trajectory demonstrates the rapid re-emergence of MT-ZVL in central China, with accelerating spatial expansion.

The transmission is influenced by complex interactions among environmental, biological, and socioeconomic factors.[30,31] Although existing ecological niche models have incorporated environmental and socioeconomic variables to predict the risk of MT-ZVL transmission,[7,8,11] their accuracy is limited by three key

factors: (i) the exclusion of essential biological factors (vectors/infectious canines)[7]; (ii) a reliance on potentially distorted socioeconomic proxies (e.g., population density, GDP), which are unreliable given China's massive internal migration during accelerated urbanization[32]; and (iii) the utilization of low-resolution human case records, limited to county-level administrative divisions.[25] Our study overcomes these limitations by integrating high-resolution environmental data with township-level case records, vector surveillance, and canine infection data, allowing more-precise risk mapping.

The model predicted continuous high-risk zones across central China's mountainous regions, with fragmented suitability in eastern/southern Gansu and northern Sichuan. These high-risk areas not only match known endemic counties,[7,8,10] but also include counties reporting few or any local MT-ZVL cases, highlighting the potential risk of emergence if infectious reservoirs emerge.[12] Field surveys revealed high *P. chinensis* densities in Jiexiu, and MT-ZVL resurgence in neighboring counties,[10] confirming the substantial transmission risk in these areas. The model also identified historically endemic areas in southern Gansu and northern Sichuan, where recent case reports (2019–2021) have remained sparse despite previous higher prevalence.[10] These regions showed a moderate but patchy risk, with suitability concentrated in key counties.[10] Compared with previous predictions based on the occurrence of either *P. chinensis* or human MT-ZVL cases, our multi-source model predicted a more spatially extensive distribution of the MT-ZVL risk.[7,8,11] The most significant increases in risk were found along the edges of the Taihang, Lvliang, and Qinling mountain ranges. This refined prediction better matched the observed case distribution, especially in central and southern Shanxi, western Hebei, Beijing, eastern Gansu, and northwestern Henan[7,10]—areas in which sporadic local human cases have emerged in recent years. Our risk mapping also revealed significant northward spread of the MT-ZVL risk into new areas—particularly northern Shanxi, Hebei, and Beijing. These findings, undetected in previous assessments, highlight the importance of the integration of up-to-date data in tracking the shifting distribution of MT-ZVL, which is potentially driven by climate change.[7,10]

The risk of MT-ZVL transmission exceeded our conventional epidemiological understanding.[7,8,11] Our modeling estimated that 46.67 million people reside in at-risk areas, spanning 291 counties/districts across 10 provinces, including 117 confirmed endemic counties and 174 newly identified at-risk counties. High- or medium-risk zones—including counties with established transmission and those with resurgence potential counties with canine leishmaniasis—spanned 210 counties. This is consistent with the results of the surveillance of infectious canine infections in potential risk zones in 2022, which identified 196 at-risk counties (125 early-resurgence counties with VL-affected canines and 71 endemic counties).[10] A persistent high burden of asymptomatic *Leishmania* infections in endemic areas confirmed the substantial at-risk populations,[1] despite rural depopulation due to urbanization. These risk stratification patterns and the quantification of at-risk populations are directly applicable to the development of health policies and the strategic allocation of interventions to mitigate the MT-ZVL burden.[25]

Risk-tiered intervention strategies are proposed. In designated high-risk zones, integrated reservoir-vector control should be prioritized through timely diagnosis and treatment of human cases, enhanced canine population management, widespread deployment of insecticide-impregnated collars on dogs, and targeted vector suppression under the One Health approach[33,34]; Medium-risk areas require establishment of containment buffers in priority townships via focal insecticide-impregnated collars deployment to prevent geographical expansion into peripheral territories; While low-risk regions necessitate enhanced surveillance of *P. chinensis* density dynamics and canine *Leishmania* seroprevalence.

This study has several limitations. First, XGBoost captures complex nonlinear associations but cannot infer causality, despite the use of SHAP for interpretability. Second, the risk maps represent theoretical suitability rather than currently active transmission foci; they provide a basis for risk stratification but require refinement with local fine-scale data and iterative updating as new surveillance data become available.[25] Third, although the base models demonstrated acceptable discriminative ability, uncertainty from these predictions was not propagated into the top-level model. This methodological limitation may result in overly narrow confidence intervals in the final risk estimates. Future studies should adopt Bayesian hierarchical approaches to fully account for uncertainty across modeling layers. Fourth, while the model performed well across ten provinces with similar mountainous topography and historical MT-ZVL endemicity, its generalizability to ecologically distinct regions remains unvalidated. Spatial cross-validation (e.g., block or distance-constrained) should be incorporated in future work to assess geographic transferability.

This study has demonstrated that MT-ZVL is a significant public health threat, showing rapid geographic expansion across the hilly regions of central China. Key findings were: (1) the endemic spread of MT-ZVL persists unabated; (2) the confirmed presence of canine reservoirs and *P. chinensis* in currently non-endemic areas indicates increasing transmission potential and the emergence of novel endemic foci; (3) the presence of *P. chinensis* in high-latitude regions may be driving the northward expansion of MT-ZVL transmission; (4) this optimized multisource data-driven machine learning model showed robust performance in

predicting MT-ZVL distribution and spread patterns, and strong performance in external validation; and (5) this study provides a precise analysis of the epidemiological situation, identifying 291 at-risk counties/districts, including 117 confirmed endemic counties and 174 newly identified at-risk counties across 10 provinces, with approximately 46.67 million residents at risk. Consequently, our study provides critical guidance for evidence-based interventions to control MT-ZVL.

Contributors

Zhengbin Zhou, Lulu Huang, Zhongqiu Li, Yuwan Hao, Yuanyuan Li, Haobo Ni, Limin Yang, Xinyi Wang, Canjun Zheng, Ying Liu, Xiaodong Tian, Wenting Wu, Qin Liu, Yi Zhang and Junhu Chen collected the data. Yanfeng Gong, Zhengbin Zhou and Shizhu Li performed the analysis. Zhengbin Zhou and Yanfeng Gong drafted the manuscript. Zhengbin Zhou, Yanfeng Gong, Shizhu Li and Qun Li conceived and designed this study. Zhengbin Zhou, Yanfeng Gong, Shang Xia, Junhu Chen, Yibiao Zhou, Xiaonong Zhou, Robert Berquist, Shizhu Li and Qun Li made critical revisions of the manuscript. Shizhu Li and Qun Li led the successful funding application. Zhengbin Zhou and Yanfeng Gong have directly accessed and verified the data. All authors approved the final manuscript and agreed to submit the manuscript.

Data sharing statement

The publicly available environmental data and field survey data collected during 2022–2024 and used in this study can be accessed directly from the databases listed in Table 1 and Supplementary Table S1. Human case data obtained from the National Notifiable Disease Surveillance System (NNDSS) are not publicly available due to public health data management regulations and privacy protection requirements. However, researchers may submit reasonable requests for data access to the corresponding author.

Editor note

The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We extend our sincere gratitude to all provincial and county-level Centers for Disease Control (CDCs) whose invaluable support made this investigation possible.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanwpc.2026.101858.

References

1 Burza S, Croft SL, Boelaert M. Leishmaniasis. *Lancet*. 2018;392(10151):951–970.
2 Lun ZR, Wu MS, Chen YF, et al. Visceral leishmaniasis in China: an endemic disease under control. *Clin Microbiol Rev*. 2015;28(4):987–1004.
3 WHO. Leishmaniasis. Geneva: World Health Organization https://www.who.int/news-room/fact-sheets/detail/leishmaniasis. Accessed January 12, 2023.
4 Casulli A. New global targets for NTDs in the WHO roadmap 2021-2030. *PLoS Negl Trop Dis*. 2021;15(5):e0009373.
5 Zhou ZB, Wang JY, Gao CH, et al. Contributions of the National Institute of Parasitic Diseases to the control of visceral leishmaniasis in China. *Adv Parasitol*. 2020;110:185–216.
6 Moirano G, Zanet S, Giorgi E, et al. Integrating environmental, entomological, animal, and human data to model the Leishmania infantum transmission risk in a newly endemic area in Northern Italy. *One Health*. 2020;10:100159.
7 Meng Z, Fan PW, Fan ZX, et al. Environmental change increases the transmission risk of visceral leishmaniasis in central China around the Taihang mountains. *Environ Health*. 2025;24(1):27.
8 Zhao Y, Jiang D, Ding F, et al. Recurrence and driving factors of visceral leishmaniasis in central China. *Int J Environ Res Public Health*. 2021;18(18):9535.
9 Yang C, Li S, Lu D, et al. Reemergence of visceral leishmaniasis in Henan Province, China. *Trop Med Infect Dis*. 2023;8(6):318.
10 Zhou Z, Luo Z, Pan G, et al. Epidemiological features and spatial-temporal clustering of visceral leishmaniasis - china, 2011-2022. *China CDC Wkly*. 2024;6(46):1201–1205.
11 Wang X, Xue J, Xia S, et al. Distribution of suitable environments for Phlebotomus chinensis as the vector for mountain-type zoonotic visceral leishmaniasis - six provinces, China. *China CDC Wkly*. 2020;2(42):815–819.
12 Liu G, Wu Y, Wang L, et al. Re-emergence of canine Leishmania infantum infection in mountain areas of Beijing. *One Health Adv*. 2023;1:11.
13 Kaempfle M, Hartmann K, Bergmann M. Treatment of Leishmania infantum infections in dogs. *Microorganisms*. 2025;13(5):1018.
14 Dantas-Torres F, Miró G, Bowman DD, Gradoni L, Otranto D. Culling dogs for zoonotic visceral leishmaniasis control: the wind of change. *Trends Parasitol*. 2019;35(2):97–101.
15 Xue J, Hu X, Hao Y, et al. Transmission risk predicting for schistosomiasis in mainland China by exploring ensemble ecological niche modeling. *Trop Med Infect Dis*. 2022;8(1):24.
16 Townsend JP, Aldstadt J. Habitat suitability mapping using logistic regression analysis of long-term bioacoustic bat survey dataset in the Cassadaga Creek watershed (USA). *Sci Total Environ*. 2023;895:165077.
17 Romero-Alvarez D, Escobar LE, Auguste AJ, Del Valle SY, Manore CA. Transmission risk of Oropouche fever across the Americas. *Infect Dis Poverty*. 2023;12(1):47.
18 Ong J, Liu X, Rajarethinam J, et al. Mapping dengue risk in Singapore using Random Forest. *PLoS Negl Trop Dis*. 2018;12(6):e0006587.
19 Ahangarcani M, Farnaghi M, Shirzadi MR, Pilesjö P, Mansourian A. Predictive risk mapping of human leptospirosis using support vector machine classification and multilayer perceptron neural network. *Geospat Health*. 2019;14(1). https://doi.org/10.4081/gh.2019.711.
20 Wang T, Ge Q, Ma T, et al. A novel method for predicting debris flow hazard: a multi-strategy fusion approach based on the light gradient boosting machine framework. *Stoch Environ Res Risk Assess*. 2025:1–24.
21 Xu N, Zhang Y, Du C, et al. Prediction of Oncomelaniahupensis distribution in association with climate change using machine learning models. *Parasit Vectors*. 2023;16(1):377.
22 Montgomery-Csobán T, Kavanagh K, Murray P, et al. Machine learning-enabled maternal risk assessment for women with pre-eclampsia (the PIERS-ML model): a modelling study. *Lancet Digit Health*. 2024;6(4):e238–e250.
23 Gong YF, Hu XK, Hao YW, et al. Projecting the proliferation risk of Oncomelaniahupensis in China driven by SSPs: a multi-scenario comparison and integrated modeling study. *Adv Clim Change Res*. 2022;13(2):258–265.
24 Sanchez MCA, Celeste BJ, Lindoso JAL, et al. Performance of rK39-based immunochromatographic rapid diagnostic test for serodiagnosis of visceral leishmaniasis using whole blood, serum and oral fluid. *PLoS One*. 2020;15(4):e0230610.
25 Lim A, Shearer FM, Sewalk K, et al. The overlapping global distribution of dengue, chikungunya, Zika and yellow fever. *Nat Commun*. 2025;16(1):3418.
26 Guan LR. Current status of kala-azar and vector control in China. *Bull World Health Organ*. 1991;69(5):595–601.
27 Ferroglio E, Maroli M, Gastaldo S, Mignone W, Rossi L. Canine leishmaniasis, Italy. *Emerg Infect Dis*. 2005;11(10):1618–1620.
28 Luo Z, Zhou Z, Hao Y, et al. Establishment of an indicator framework for the transmission risk of the mountain-type zoonotic visceral leishmaniasis based on the Delphi-entropy weight method. *Infect Dis Poverty*. 2022;11(1):122.
29 Quinnell RJ, Carson C, Reithinger R, Garcez LM, Courtenay O. Evaluation of rK39 rapid diagnostic tests for canine visceral leishmaniasis: longitudinal study and meta-analysis. *PLoS Negl Trop Dis*. 2013;7(1):e1992.

30 Diaz-Saez V, Corpas-Lopez V, Merino-Espinosa G, Morillas-Mancilla MJ, Abattouy N, Martin-Sanchez J. Seasonal dynamics of phlebotomine sand flies and autochthonous transmission of leishmania infantum in high-altitude ecosystems in Southern Spain. *Acta Trop*. 2021;213:105749.
31 Leal Filho W, Nagy GJ, Gbaguidi GJ, et al. The role of climatic changes in the emergence and re-emergence of infectious diseases: bibliometric analysis and literature-supported studies on zoonoses. *One Health Outlook*. 2025;7(1):12.
32 Mellander C, Lobo J, Stolarick K, et al. Night-time light data: a good proxy measure for economic activity? *PLoS One*. 2015;10(10):e0139779.
33 Yimam Y, Mohebali M. Effectiveness of insecticide-impregnated dog collars in reducing incidence rate of canine visceral leishmaniasis: a systematic review and meta-analysis. *PLoS One*. 2020;15(9):e0238601.
34 Zhang XX, Guo XK, Zhou XN. One health: a key element in the WHO pandemic agreement. *Lancet*. 2025;405(10496):2197–2198.