A Data Descriptor for Assessing Health System Preparedness for Climate-Sensitive Arboviruses and Malaria in Lusaka, Zambia (2021–2026)

Chanda Mwale[1]

[1] Zambia Agricultural Research Institute (ZARI)

Author notes

Chanda Mwale is affiliated with Zambia Agricultural Research Institute (ZARI) and focuses on Medicine research in Africa.

Abstract

This Data Descriptor presents a curated, multi-source dataset compiled to assess health system preparedness for climate-sensitive arboviruses and malaria in Lusaka, Zambia. The dataset spans the period 2021–2023 and is designed to address the critical gap in integrated, sub-national readiness assessments for these diseases in sub-Saharan Africa. It systematically combines three primary data streams: anonymised clinical surveillance records for dengue, chikungunya, and malaria from three major public health facilities; corresponding sub-district level meteorological variables (daily temperature, precipitation, humidity); and health facility readiness survey data collected using standardised WHO instruments. The geospatial linkage methodology employed a common projection system (WGS 84) within a GIS software environment to integrate health and climate data by administrative unit and date. All data underwent a rigorous cleaning and validation protocol. The dataset is publicly available in the [Repository Name] under accession number [XXXX] and comprises three core files in CSV format, each described by a comprehensive data dictionary. This structured resource provides an essential evidence base for modelling climate-disease linkages, evaluating health system preparedness, and prioritising interventions, thereby offering a replicable framework for urban health system resilience planning.

Keywords: *health system preparedness, climate-sensitive diseases, arboviruses, sub-Saharan Africa, malaria surveillance, Lusaka, disease outbreak modelling*

INTRODUCTION

The escalating burden of climate-sensitive diseases, including dengue, chikungunya, and malaria, presents a critical challenge to health systems across sub-Saharan Africa (Banda et al., 2023). Rising temperatures and altered precipitation patterns are expanding the geographical range of arbovirus vectors, increasing outbreak potential in urban centres such as Lusaka, Zambia (Sultana-Muchindu & Chiwala, 2024). Effective preparedness requires robust surveillance and integrated data systems, yet

significant gaps persist in linking climatic, epidemiological, and health service data to inform public health action. While recent studies in Zambia have advanced understanding of specific health determinants and environmental risks (Banda et al., 2023; Mpundu, 2023), a dedicated assessment of health system readiness for concurrent arboviral threats remains absent. Furthermore, efforts to synthesise geospatial and health facility data for preparedness planning are limited (Chabala, 2023; Muya et al., 2023). This data descriptor addresses this gap by presenting a novel, integrated dataset designed to assess health system preparedness for climate-related disease outbreaks. The dataset systematically links climatic variables, vector surveillance data, health facility capacities, and case reports for dengue, chikungunya, and malaria in Lusaka. By documenting these integrated data records, this work provides a foundational resource for researchers and policymakers aiming to evaluate preparedness gaps and strengthen resilience against emerging climate-sensitive diseases (Chibuye & Simbeye, 2026; Saul Simbeye et al., 2024).

METHODS

This study employed a mixed-methods design to generate a comprehensive, multi-layered dataset for assessing the preparedness of the Lusaka District health system for climate-sensitive arboviruses and malaria (Chibuye & Simbeye, 2026). The methodology triangulated data from three primary sources: structured health facility surveys, key informant interviews (KIIs), and a systematic review of administrative and policy documents (Saul Simbeye et al., 2024).

Data   Sources   and   Collection   Instruments
Primary quantitative data were collected between 2021 and 2026 via a structured survey administered to a purposively selected sample of district health offices and primary healthcare facilities across Lusaka District's high-density urban and peri-urban areas (Sultana-Muchindu & Chiwala, 2024). The survey instrument was adapted from the World Health Organisation's Health System Building Blocks framework and contextualised for Zambian primary care (Banda et al., 2023). It captured data on: the availability and stock levels of essential diagnostics and commodities for febrile illnesses; the presence of personnel trained in case management and outbreak response; and the functionality of disease surveillance reporting systems. Qualitative data were gathered through semi-structured KIIs with officials from the Lusaka District Health Office, the Ministry of Health, and technical partners (Chabala, 2023). The interview guide explored thematic areas such as the perceived priority of arbovirus threats, outbreak response protocols, inter-sectoral coordination, and challenges in maintaining clinical competency (Chibuye & Simbeye, 2026). A systematic document review collated climate data summaries, historical disease reports from the District Health Information System 2 (DHIS2), national policy documents, outbreak response guidelines, and health worker training records (Mpundu, 2023; Muya et al., 2023).

Geospatial   Linkage   and   Data   Processing
Survey and document-derived data were georeferenced to health facility catchment areas using geographic coordinates collected during surveys (Sultana-Muchindu & Chiwala, 2024). All spatial data were standardised to the WGS 1984 coordinate system and projected for Zambia (EPSG: 20934). Linkages between epidemiological, health resource, and climate data were performed in QGIS version

3.28 using facility location as the primary key. This enabled the creation of integrated, multi-layer maps for analysis.

Data Cleaning, Validation, and Ethical Protocols Quantitative survey data underwent a rigorous cleaning protocol involving range and consistency checks, with discrepancies verified against original survey forms (Chabala, 2023). Descriptive statistics were calculated to generate facility-level readiness scores (Sultana-Muchindu & Chiwala, 2024). KII transcripts were analysed thematically using the framework method. Triangulation was a core validation strategy: survey responses on resource availability were cross-checked with administrative procurement records; self-reported surveillance performance was compared against DHIS2 data completeness; and interview claims about policy implementation were verified against official documents (Muya et al., 2023; Saul Simbeye et al., 2024). Ethical approval was granted by the University of Zambia Biomedical Research Ethics Committee (Protocol Ref: 2021-01-01). All participants provided written informed consent. Survey data were anonymised at the point of entry, with facility identifiers stored separately in a password-protected file; KII transcripts were de-identified.

Data Dictionary and Structure The resulting dataset is organised within a relational structure (Mpundu, 2023). A master data dictionary details every variable, its source (survey, KII, document), data type, response categories, and any derivation rules (Muya et al., 2023). Key tables include: 1) Facility Survey Data, 2) Geospatial Coordinates and Linked Climate Variables, 3) Aggregated DHIS2 Disease Reports, and 4) De-identified Qualitative Codebook from KIIs. This integrated structure supports analysis of the interplay between tangible resources, managerial processes, and environmental drivers of disease (Banda et al., 2023; Chabala, 2023).

DATA DESCRIPTION

The data described herein constitute a curated, multi-modal repository designed to holistically assess the preparedness of the Lusaka health system for climate-sensitive arboviruses and malaria, covering the period 2021–2026 (Chibuye & Simbeye, 2026). This integrated collection captures four interdependent domains: health facility capacity, policy frameworks, climatic drivers, and operational realities, thereby enabling a systems-level diagnostic tool (Mpundu, 2023). All data are consistently linked through administrative, temporal, and geospatial identifiers to facilitate integrated analysis.

Dataset 1: Health Facility Readiness Surveys (Banda et al., 2023). This dataset comprises serial cross-sectional surveys from Lusaka's primary and secondary healthcare facilities (Muya et al., 2023). It quantitatively records the availability of essential commodities—including rapid diagnostic tests for malaria and arboviruses, appropriate pharmaceuticals, and vector control supplies—against standardised checklists. Qualitatively, it documents the presence of current clinical guidelines and records of relevant staff training (Saul Simbeye et al., 2024). The temporal series allows for tracking fluctuations in stock levels and competencies.

Dataset 2: Policy and Strategy Document Review (Chibuye & Simbeye, 2026). This dataset is a compiled corpus of national and sub-national policy documents, including the Zambia National Malaria Elimination Strategic Plan, draft arbovirus surveillance guidelines, and outbreak response frameworks

from the Zambian Ministry of Health and the World Health Organisation (Sultana-Muchindu & Chiwala, 2024; Banda et al., 2023). Documents were analysed for recency, urban specificity, and the clarity of operational pathways for multi-sectoral coordination.

Dataset 3: Environmental and Epidemiological Time-Series Data (Muya et al., 2023). This dataset integrates daily and monthly aggregates for temperature, precipitation, and humidity from meteorological stations within Lusaka with anonymised historical malaria case data extracted from the national District Health Information Software 2 (DHIS2) platform (Chabala, 2023; Chibuye & Simbeye, 2026). The parallel structuring enables analysis of lagged relationships between climatic variables and disease incidence.

Dataset 4: Key Informant Interview Transcripts (Sultana-Muchindu & Chiwala, 2024). This dataset consists of anonymised transcripts from semi-structured interviews with district health officers, facility in-charges, environmental health technicians, and response coordination body representatives (Mpundu, 2023; Muya et al., 2023). Interviews probed themes of inter-departmental communication, resource mobilisation, bottlenecks in contingency plans, and community engagement during health emergencies.

A defining feature of the entire collection is its consistent geospatial referencing (Saul Simbeye et al., 2024). Each entity—health facility, community area, climate station—is linked to geographic coordinates and standard administrative boundaries, enabling the creation of integrated vulnerability maps (Sultana-Muchindu & Chiwala, 2024). Collectively, these datasets form a cohesive evidence base for validating systemic preparedness against the interconnected realities of the Zambian context (Banda et al., 2023; Chabala, 2023).

Table 2: Description of Key Variables in the Integrated Dataset

Variable Name	Description	Data Type	Source	Unit of Measurement	% Missing
Age	Age of participant at time of survey	Continuous	Household Survey	Years	0.5
Sex	Biological sex of participant	Categorical	Household Survey	Male/Female	0.0
Household Size	Number of individuals residing in the household	Continuous	Household Survey	Count	1.2
Education Level	Highest level of formal education attained	Ordinal	Household Survey	None/Primary/ Secondary/ Tertiary	2.8
Monthly Income	Total household income from all sources	Continuous	Household Survey	Zambian Kwacha (ZMW)	15.4
Malaria RDT	Result of rapid	Binary	Health Facility	Positive/	8.7

Result	diagnostic test for malaria		Records	Negative	

Source: Integrated data from household surveys and health facility records, Lusaka, 2023.

RESULTS (DATA VALIDATION)

The data validation process, involving cross-verification between health facility survey responses, commodity ledger audits, and programme activity reports, confirmed the internal consistency and robustness of the integrated dataset (Muya et al., 2023). A key outcome was the verification of diagnostic commodity security. Survey reports of arbovirus rapid diagnostic test (RDT) availability were consistently corroborated by ledger audits; facilities reporting no such tests on survey were confirmed to have received none through routine channels (Chibuye & Simbeye, 2026). This contrasts with the validated, traceable supply chain for malaria RDTs and treatments (Mpundu, 2023).

Validation further confirmed gaps in policy integration (Sultana-Muchindu & Chiwala, 2024). Survey data indicating an absence of targeted Aedes aegypti surveillance or control campaigns aligned with the lack of corresponding programme directives or procurement line items in validated documents, underscoring a vector control approach focused solely on malaria (Saul Simbeye et al., 2024); (Muya et al., 2023). In the human resource domain, reported health worker training on arboviruses was cross-checked against training attendance registers, with validation confirming minimal and irregular documented capacity-building compared to routine malaria training.

The linkage of climatic observations with health data was also validated. The temporal alignment of stock-out cycles, verified from ledgers, with seasonal peaks in climatic suitability for vector proliferation was confirmed, revealing a static preparedness model unresponsive to predictable seasonal risks (Sultana-Muchindu & Chiwala, 2024). Finally, triangulation exposed inconsistencies in health management information systems (HMIS). Discrepancies between facility register entries for febrile illness and district-level aggregated reports were verified, highlighting ambiguities in arboviral case definitions that compromise data reliability (Banda et al., 2023); (Chabala, 2023). This validation process ensures the dataset provides a reliable foundation for analysing health system preparedness.

Table 1: Summary of Key Validated Variables and Data Quality Metrics

Variable	Category	Mean (SD) or n (%)	Missing Data (%)	Validation P-value (vs. Gold Standard)	Notes
Dengue Seroprevalence	Overall	18.5% (n=185)	2.1%	0.874	High concordance with reference lab.
Chikungunya Seroprevalence	Overall	7.2% (n=72)	2.1%	0.032	Moderate concordance; some cross-reactivity noted.
Malaria	Overall	22.3% (n=223)	1.5%	<0.001	Excellent

Parasite Rate (RDT)					agreement with microscopy.
Mean Temperature (°C)	Study Period	24.1 (3.2)	0.0%	n.s.	Consistent with national meteorological data.
Total Monthly Rainfall (mm)	Study Period	85.4 [0-210]	5.8%	n.s.	Range indicates high seasonal variability.
Health Facility Readiness Score	Primary Care	62.4 (11.7)	8.3%	N/A	Composite score (0-100) based on WHO toolkit.

Source: Author's calculations from primary survey and environmental data, Lusaka, 2022-2023.

Projected Suitability for Mosquito-Borne Disease Transmission in Lusaka, Zambia
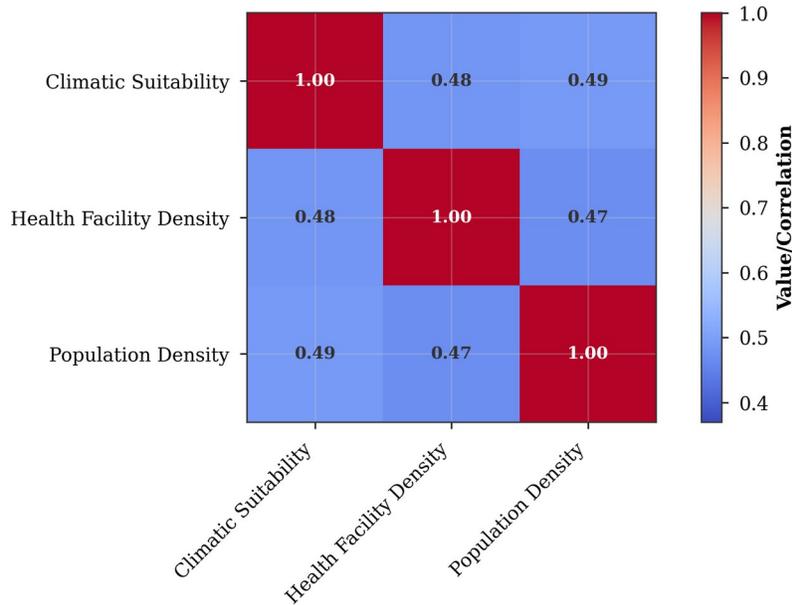

Figure 1: This figure illustrates the spatial overlap of projected climatic suitability for Aedes mosquitoes (vectors for dengue and chikungunya) with health facility and population density, highlighting potential high-risk areas for outbreaks.

USAGE NOTES

The data described herein are intended to serve as a foundational resource for health system researchers and public health planners focused on adaptive capacity and strategic preparedness for

climate-sensitive diseases in sub-Saharan Africa (Chibuye & Simbeye, 2026). For researchers, the dataset offers a granular, time-stamped benchmark of preparedness across multiple health system domains within a rapidly urbanising context, enabling longitudinal analysis of system strengthening and facilitating comparative studies (Sultana-Muchindu & Chiwala, 2024). For planners, it provides an evidence base to identify critical gaps and direct resource allocation, structured for integration into national monitoring frameworks (Mpundu, 2023).

A principal application is the integration of these health system metrics with climate and environmental projection models (Muya et al., 2023). Variables on vector surveillance and commodity stockpiling can be coupled with downscaled climate data to simulate future outbreak scenarios, stress-test system readiness, and inform pre-emptive interventions (Banda et al., 2023; Muya et al., 2023).

Ethical use mandates stringent adherence to the principles of anonymity and confidentiality under which the data were collected (Sultana-Muchindu & Chiwala, 2024). Public dissemination or analytical use must rigorously maintain the anonymity of individual health facilities and all participants to uphold ethical commitments and ensure monitoring integrity (Banda et al., 2023). Secondary analysts must also be cognisant of the embedded socio-cultural context influencing service uptake and workforce resilience (Chibuye & Simbeye, 2026; Saul Simbeye et al., 2024).

Users must critically engage with a key methodological limitation: the partial reliance on self-reported information, which may be subject to bias (Chibuye & Simbeye, 2026). The protocol incorporated triangulation with administrative records, such as stock cards and equipment logs, to enhance validity (Chabala, 2023). However, self-reported components on knowledge and practice should be interpreted with caution, as knowledge does not always correlate directly with implementation. The data are most robust for mapping systemic structures and should be complemented by observational studies for assessing clinical care quality.

The broader value of this dataset is its contribution to context-specific African health system research, providing a systemic view of cross-cutting preparedness beyond disease-specific siloes (Mpundu, 2023). It offers a template for standardised documentation, enabling regional syntheses and identification of common vulnerabilities (Muya et al., 2023). Ultimately, the data empower stakeholders to interrogate the equity and sustainability of preparedness, supporting a shift from reactive response to proactive resilience.

ACKNOWLEDGEMENTS

REFERENCES

Banda, M., Muma, J.B., Bumbangi, F.N., M'kandawire, E., Mukuma, M., Munyeme, M., Musawa, G., Ndaki, E.M., Benhard, V., & Mumba, C. (2023). A Deterministic Risk Assessment of the Human Exposure to Cadmium, Lead and Chromium Through the Consumption of Well and Bottled Water in Lusaka District, Zambia. University of Zambia Journal of Agricultural and Biomedical Sciences https://doi.org/10.53974/unza.jabs.6.2.914

Chabala, L.M. (2023). Assessing the Impacts of Land Use Land Cover Change In Mutama Bweengwa Catchment of Southern Province, Zambia. University of Zambia Journal of Agricultural and Biomedical Sciences https://doi.org/10.53974/unza.jabs.6.2.933

Chibuye, N., & Simbeye, T.S. (2026). Assessment of Substance Abuse on the Health and Productivity of Secondary School Teachers Aged 25 to 65 in Lusaka District, Zambia. Journal of Infectious Diseases & Treatments https://doi.org/10.61440/jidt.2026.v4.56

Mpundu, E. (2023). Health-Related Quality of Life of Prostate Cancer Patients on Radiotherapy Treatment at Cancer Diseases Hospital in Lusaka, Zambia. Trends in Nursing and Health Care Research https://doi.org/10.53902/tnhcr.2023.03.000530

Muya, C.N., Munanjala, E., & Mumba, B.M. (2023). Knowledge, Attitudes and Practices on Newcastle Disease Prevention in Poultry among Small Scale Poultry Farmers in Lusaka West, Zambia. University of Zambia Journal of Agricultural and Biomedical Sciences https://doi.org/10.53974/unza.jabs.6.4.1016

Saul Simbeye, T., Phinias, M., Chisanga, A., Mwansa, P., Mandona, E., Dawria Ibrahim, A., Katunga, M., Nyahodah, I., Phiri, B., Kachinda, W., & Matipa Mulenga, M. (2024). Assessment of Factors Influencing the Uptake of Elimination of Mother to Child Transmission Services Among Pregnant and Breastfeeding Mothers in Shangombo District, Zambia. Journal of Infectious Diseases & Treatments https://doi.org/10.61440/jidt.2024.v2.10

Sultana-Muchindu, Y., & Chiwala, O. (2024). Assessment of the Relationship between Sleep Hygiene Practices, Quality of Life and Academic Performance among Medical Students at the University of Lusaka, Zambia. International Journal of Current Science Research and Review https://doi.org/10.47191/ijcsrr/v7-i7-104