



OPEN

Geo-epidemiology of malaria incidence in the Vhembe District to guide targeted elimination strategies, South-Africa, 2015–2018: a local resurgence

Sokhna Dieng^{1✉}, Temitope Christina Adebayo-Ojo², Taneshka Kruger³, Megan Riddin³, Helene Trehard¹, Serena Tumelero¹, Marc-Karim Bendiane¹, Christiaan de Jager³, Sean Patrick³, Riana Bornman³ & Jean Gaudart⁴

In South Africa, the population at risk of malaria is 10% (around six million inhabitants) and concern only three provinces of which Limpopo Province is the most affected, particularly in Vhembe District. As the elimination approaches, a finer scale analysis is needed to accelerate the results. Therefore, in the process of refining local malaria control and elimination strategies, the aim of this study was to identify and describe malaria incidence patterns at the locality scale in the Vhembe District, Limpopo Province, South Africa. The study area comprised 474 localities in Vhembe District for which smoothed malaria incidence curve were fitted with functional data method based on their weekly observed malaria incidence from July 2015 to June 2018. Then, hierarchical clustering algorithm was carried out considering different distances to classify the 474 smoothed malaria incidence curves. Thereafter, validity indices were used to determine the number of malaria incidence patterns. The cumulative malaria incidence of the study area was 4.1 cases/1000 person-years. Four distinct patterns of malaria incidence were identified: high, intermediate, low and very low with varying characteristics. Malaria incidence increased across transmission seasons and patterns. The localities in the two highest incidence patterns were mainly located around farms, and along the rivers. Some unusual malaria phenomena in Vhembe District were also highlighted as resurgence. Four distinct malaria incidence patterns were found in Vhembe District with varying characteristics. Findings show also unusual malaria phenomena in Vhembe District that hinder malaria elimination in South Africa. Assessing the factors associated with these unusual malaria phenome would be helpful on building innovative strategies that lead South Africa on malaria elimination.

Abbreviation

IQR Interquartile range

In 2021, the World Health Organization (WHO) estimated 247 million malaria cases and 625,000 malaria deaths in 84 malaria-endemic countries. The WHO African Region represented about 95% of malaria cases and deaths¹. South Africa is one of the countries in the WHO African Region² affected by malaria and progressing towards elimination. The total population at risk of malaria in South Africa represents 10% or about six million inhabitants^{3,4} and is spread over only three provinces (out of a total of nine) endemic for malaria: Limpopo, Mpumalanga and KwaZulu-Natal^{5–7}. In addition, malaria is seasonal, with a low transmission (winter) season between June and August and a high transmission (summer) season between September and May of the following year^{5,8}.

¹Aix Marseille Univ, IRD, INSERM, ISSPAM, SESSTIM, 13005 Marseille, France. ²Swiss Tropical and Public Health Institute, Basel, Switzerland. ³School of Health Systems and Public Health (SHSPH), University of Pretoria Institute for Sustainable Malaria Control (UP ISMC), University of Pretoria, Pretoria, South Africa. ⁴Aix Marseille Univ, IRD, INSERM, ISSPAM, SESSTIM, APHM, Hop. La Timone, BioSTIC, Biostatistic & ICT, 13005 Marseille, France. ✉email: sdieng.stat@gmail.com

From 2015 to 2019, South Africa recorded between 10,000 and 30,000 cases of malaria per year, and the national Department of Health plans to eliminate malaria (i.e. eliminate local transmission) by 2023⁴. However, there are challenges to achieving this goal⁴. Some researchers have differing views on the feasibility of malaria elimination, but generally agree that elimination in South Africa requires, among other things, good surveillance systems, the availability and the use of all effective intervention tools^{9,10}. Furthermore, as countries, sub-national areas and communities approach the elimination phase of malaria, they are at different epidemiological stages, and therefore, intervention strategies need to be tailored to different regions within a country¹¹. In addition, towards the elimination stage, transmission is generally low and spatially heterogeneous, thus requiring malaria analysis at the local, sub-district or village level^{11,12}. Malaria transmission is heterogeneous across South Africa's endemic provinces as showed in several studies^{8,13,14}, and progress towards elimination has been similarly variable. Strategies should be tailored to the conditions of each province. Stratification of malaria intensity level is part of the WHO recommendations to accelerate the elimination process^{11,15}. This stratification involves classifying geographical units according to local malaria transmission¹¹. The aim of this approach is to group together geographical units that have not only similar intensity levels but also similar temporal dynamics.

Therefore, the functional approach^{16,17} based on the functional data method^{18–20}, was adopted first for modelling the temporal evolution or dynamics of malaria incidence intensity, then, grouping them by similar intensity levels and temporal behaviour (dynamic) over time. The use of functional data method in public health and biomedical fields is discussed in the literature²¹. The principle of the functional data methodology is that observations (at a regular time lag or not) of a set of geographical units time series (country, health post, village, etc.) can be modelled by regularised smooth functions (called functional data or smooth curves) that are estimated with a single smoothing parameter determined from the sample. This allows the understanding of their underlying dynamics and allows the comparison of their trends to each other over time possible, because they will be analysed in the same scale¹⁸. Indeed, there was no automatic and objective way to compare thresholds and intensity levels between geographical units as shown in several epidemic studies^{22–25}. In addition, this method is very useful in the context of the increasing use of sophisticated tools to collect and store large amounts of health data such as the DHIS2²⁶. Thereafter, clustering algorithms²⁷ can be applied on these functional data (malaria smoothed curves) to identify malaria incidence patterns according to the characteristics of their dynamics (local temporal trends). This would help to guide the development and implementation of targeted control strategies in the local context.

This functional approach consisting of the of the functional data method for modelling malaria evolution of geographical units and to apply clustering algorithm to them for identifying malaria patterns, was first developed to stratify malaria incidence level, and to identify malaria patterns at a village scale in Senegal¹⁶. Then, another study adapted this to stratify and identify malaria patterns at a village scale in Myanmar¹⁷. This suggests that this approach is applicable in the context of other endemic countries, such as South Africa.

In South Africa, the Limpopo province is the most affected²⁸, particularly in the Vhembe District, where the malaria incidence in 2018 was 3.8 cases/1000 person-years²⁹. In the literature, no studies have been conducted on malaria incidence at the locality level, the smallest geographical scale in South Africa, due in part to the difficulty of population availability at this scale³⁰. As malaria elimination is approached, a finer scale analysis is needed to accelerate the results. Therefore, in the process of refining local malaria control and elimination strategies, the aim of this study was to identify and to describe the temporal dynamics of malaria incidence patterns at the locality level in the Vhembe District, Limpopo Province, South Africa.

Methods

Study area and dataset. This study was conducted in 474 localities in the Vhembe District located in the Limpopo Province of northern South Africa (Fig. 1). These localities are scattered across the four local municipalities: Musina, Thulamela, Makhado and Collins Chabane with a total population of 1,294,722 from the last national census in 2011.

The period of this study was from July 2015 to June 2018 corresponding to three malaria seasons in South Africa: 2015/2016, 2016/2017 and 2017/2018. In this study it was essential to have for each locality, in addition to the malaria cases, the population to identify the incidence patterns and the geographical coordinates to locate them on the study site.

Malaria cases. Confirmed daily malaria cases were collected by the Limpopo Provincial Malaria Programme and stored in their information system³¹. A locality is present if at least one of its inhabitants has been diagnosed with malaria during the study period.

Population. To the best of our knowledge, all the studies done on malaria incidence dynamics in South Africa were at municipality, district or province scale, where official population data is fully available at these scales. However, this is not the case at the locality scale, which is the smallest geographical scale. The South African Department of Statistics³² had population data of some localities from the last census in 2011, but not for all localities. To complete them, others data sources³³ and field investigations were used and combined with calculations based on known methods^{32,34} for all the years of the study period.

The starting point was the populations from the last national census in 2011. For each locality the information from the home municipality is provided. This was used to make a projection for each locality from 2015 to 2018 with the population growth rate of its belonging municipality corresponding to an average annual geometric variation rate^{32,34}. This rate of change per municipality RateMun_ (2001–2011) was based on the last two national censuses of 2001 and 2011.

The geographical location of the study localities in the Vhembe District of Limpopo Province in South Africa.

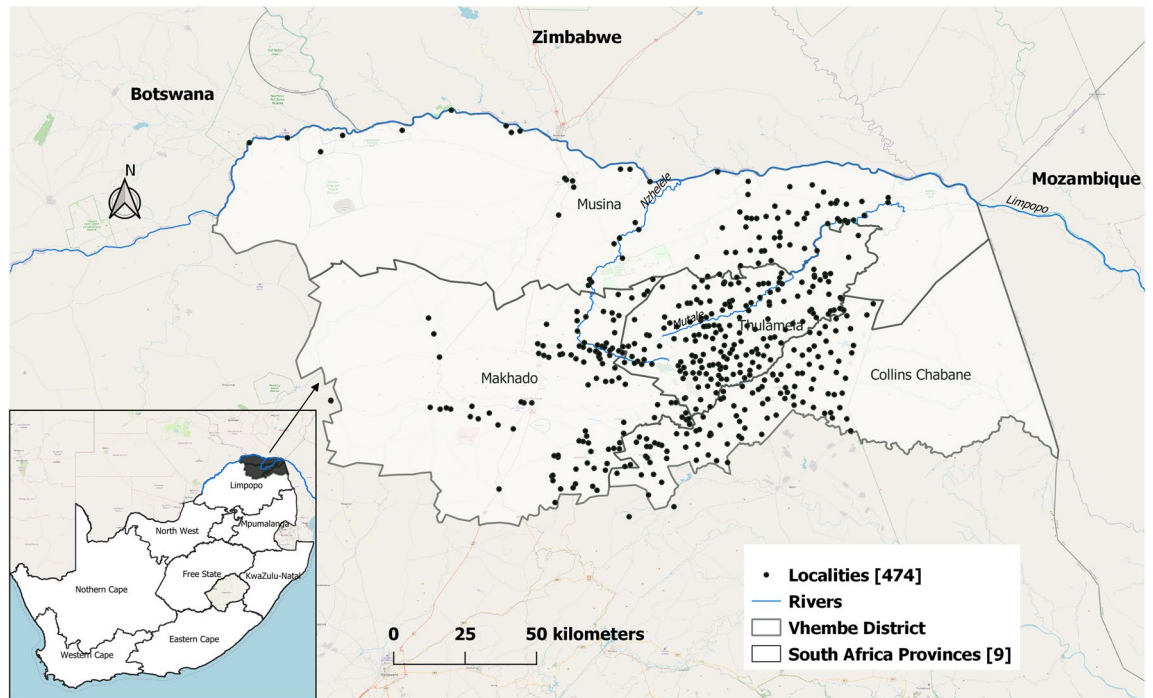


Figure 1. The study localities across Vhembe District in Limpopo Province, South Africa. Maps were produced by the authors using QGIS software (version 3.10.1., QGIS Development Team 2020. Open-Source Geospatial Foundation Project), OpenStreetMap and Shapefiles downloaded from websites^{44,45}.

$$Pop_{i,year} = Pop_{2011,i,year} * (1 + RateMun_{2001-2011i})^{year-2011}$$

Geographical coordinates. The missing geographical coordinates for some localities were obtained from the “geoview info” website³⁵, and confirmed by field investigations.

Statistical analysis. This analysis was based on an approach well described in^{16,17}. Firstly, the weekly malaria incidence time series of each locality was determined over the study period. In order to have the main trend of their evolution with reducing variability, regularized smooth function (smooth curve or functional data) was fitted from each of them with the functional data method^{18–20}.

Secondly, in order to identify patterns, hierarchical clustering algorithm with Ward’s method^{36,37} was performed on these 474 smooth functions through testing different measures of dissimilarity for having the most adaptive among them with the data. These measures of dissimilarity were the combination of temporal correlation and each of among Euclidean distance, Dynamic time warping and L2-metric in different levels of contribution^{38–41}.

Lastly, for identifying the number of patterns, Principal Components Analysis (PCA)⁴² was performed on some validity indices (Dunn, Connectivity, Silhouette width and percentage of inertia R2)^{41,43} assessed over the clustering results done with each measure of dissimilarity by considering different potential number of patterns. The final number of patterns was the potential number of models from which a clustering result had obtained the best indices: great DUNN index, small Connectivity index, Silhouette width and percentage of inertia R2 close to 1.

Statistical analyses were performed using R software (version 4.1.3., R Core Team 2020. R Foundation for Statistical Computing, Vienna, Austria). Maps were produced by the authors using QGIS software (version 3.10.1., QGIS Development Team 2020. Open-Source Geospatial Foundation Project), OpenStreetMap and Shapefiles downloaded from websites^{44,45}.

Results

Characteristics of Vhembe localities. The cumulative malaria incidence rate over the 474 localities in Vhembe was 4.1 cases/1000 person-years during the study period, from July 2015 to June 2018 (Table 1). At a localities scale, the median population was 1394 (IQR = [473; 2961]) and the median malaria incidence rate was 4 cases/1000 person-years (IQR = [1.4; 10.8]).

Over the malaria transmission season, there was an increase in both cumulative and locality malaria incidences (Table 1).

Characteristics ^a	Season 2015/2016	Season 2016/2017	Season 2017/2018	Study period
Cumulative incidence	1.0	4.0	7.4	4.1
Locality incidence	0.0 (0.0, 1.3)	3.3 (0.8, 9.4)	6.7 (2.0, 21.2)	4.0 (1.4, 10.8)

Table 1. Epidemiological characteristics of localities by seasons. ^aIncidence in cases/1000 person-years; median and interquartile range.

Main locality malaria incidence trends. From July 2015 to June 2018, time series of weekly malaria incidence for each of the 474 localities was determined (Fig. 2, panel A). A substantial variability in malaria incidence was observed among localities, making it difficult to describe their main trends. To overcome this, a smoothing was carried out on these incidences with the method of functional data as described in the methodology section. All locality time series were smoothed on the same scale with $K = 88$ base functions and $\lambda = 144$ as the smoothing parameter resulting in one smooth function (functional data or smooth curve) for each of the 474 localities in Vhembe District from July 2015 to June 2018 (Fig. 2, panel B). Thus, seasonal peaks were observed across high transmission period September to May with overall intensities increasing over time (Fig. 2, panel B). In addition, an unexpected increase was observed during the low transmission period June to August 2017.

Identification and description of malaria incidence patterns. Malaria incidence patterns were obtained after performing the hierarchical clustering procedure described in the method section on these 474 smooth curves. Using dynamic time warping dissimilarity while considering mainly the temporal correlation in the hierarchical clustering algorithm and choosing four classes, gave a simultaneously better Dunn index, width

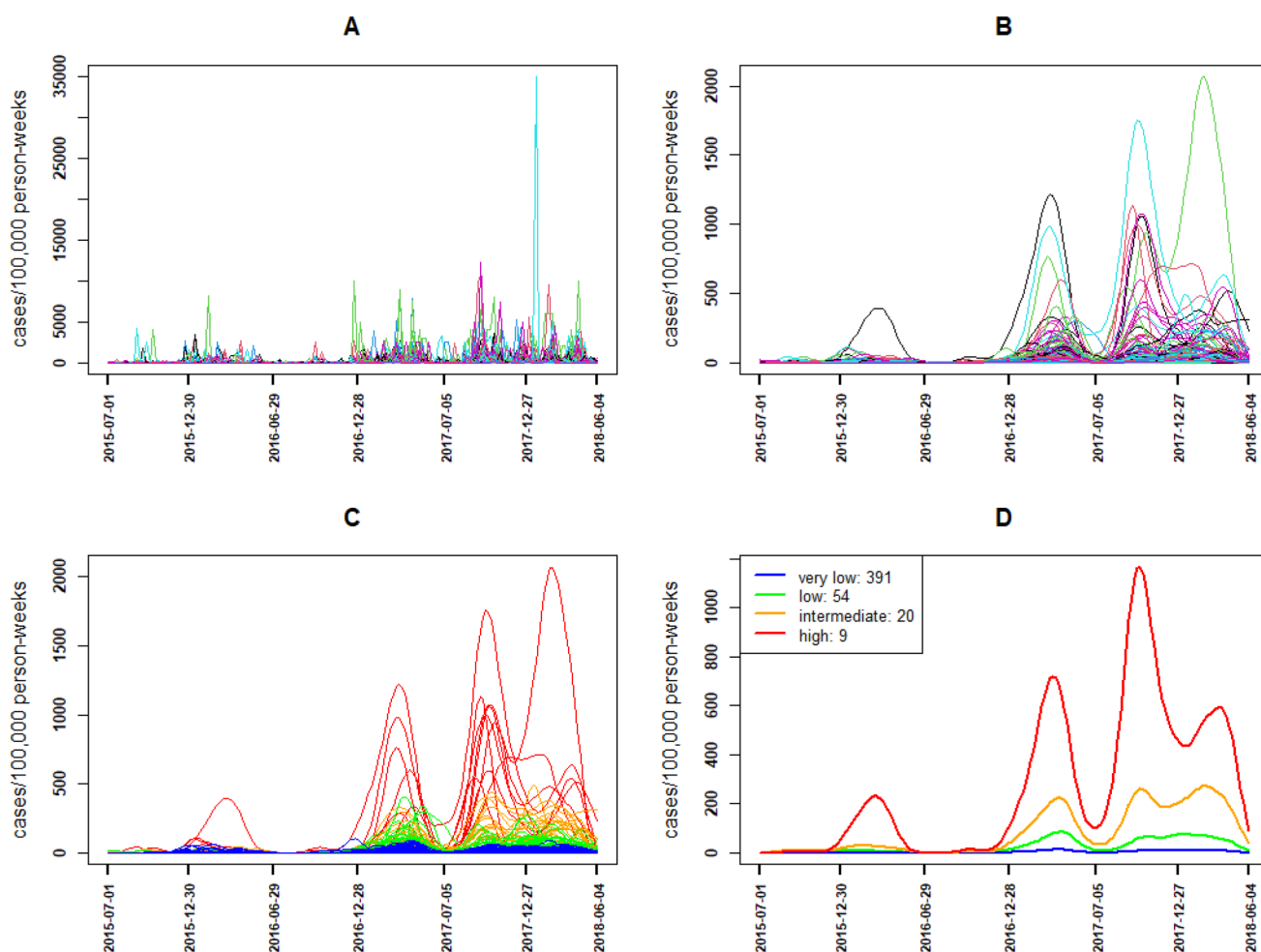


Figure 2. Smoothing and Identification of malaria incidence patterns from July 2015 to June 2018. Observed incidence localities (A), smoothed incidence localities (B), smoothed incidence localities of patterns (C) and smoothed cumulative incidence by pattern (D): very low (blue), low (green), Intermediate (orange) and high (red).

Silhouette, Connectivity index and percentage of inertia regarding PCA results. Thus, four malaria incidence patterns were obtained: very low, low, intermediate and high pattern (Fig. 2, panel C and D; 3). It is observed that localities in the same pattern had the same temporal dynamic (Fig. 3).

The very low pattern included 391 localities, the majority of which were south of the Vhembe District, between the municipalities of Makhado and Thulamela (Fig. 4). Their incidences were the lowest in the study site (Figs. 2, 3). The cumulative malaria incidence observed in this pattern during these three seasons was 2.4 cases/1000 person-years (Table 2). The median incidence over the localities in this pattern was 3 cases/1000 person-years (1.1, 6.2) (Table 2). Their seasonal peaks were more important during the end of the second malaria season 2016/2017 but their epidemic period was wider during season 2017/2018.

The low pattern included 54 localities, the majority of which were between the north and south of the Vhembe District, in the municipality of Thulamela and Musina (Fig. 4). Their incidences were the lowest in the study site after the very low profile (Figs. 2, 3). The cumulative malaria incidence observed in this pattern during these three malaria seasons was 16.2 cases/1000 person-years (Table 2). The median locality incidence was 21.7 (14.7, 33.3) cases/1000 person-years (Table 2). Their seasonal peaks were also more important during the end of the second malaria season 2016/2017, but their epidemic period was wider during season 2017/2018.

The intermediate pattern comprised 20 localities, the majority of which were in the local municipality of Musina, very few of the localities in this pattern were present in the municipalities of Makhado and Thulamela (Fig. 4). Their incidences were the second highest after those observed in the high pattern (Figs. 2, 3). The cumulative malaria incidence observed in this pattern during these three seasons was 54.1 cases/1000 person-years (Table 2) while the median locality incidence was 57.3 (40.6, 71.0) cases/1000 person-years (Table 2). Their seasonal peaks were also important during the end of the third malaria season 2017/2018 and their epidemic period was wider during this season.

The high pattern included nine localities located north in the local municipality of Musina close to river systems (Fig. 4). The incidences in these localities were overall the highest during the study period (Figs. 2, 3). The cumulative malaria incidence observed in this pattern during these three seasons was 181.2 cases/1000 person-years (Table 2). The median locality incidence was 211.3 (144.8, 237.2) cases/1000 person-years (Table 2).

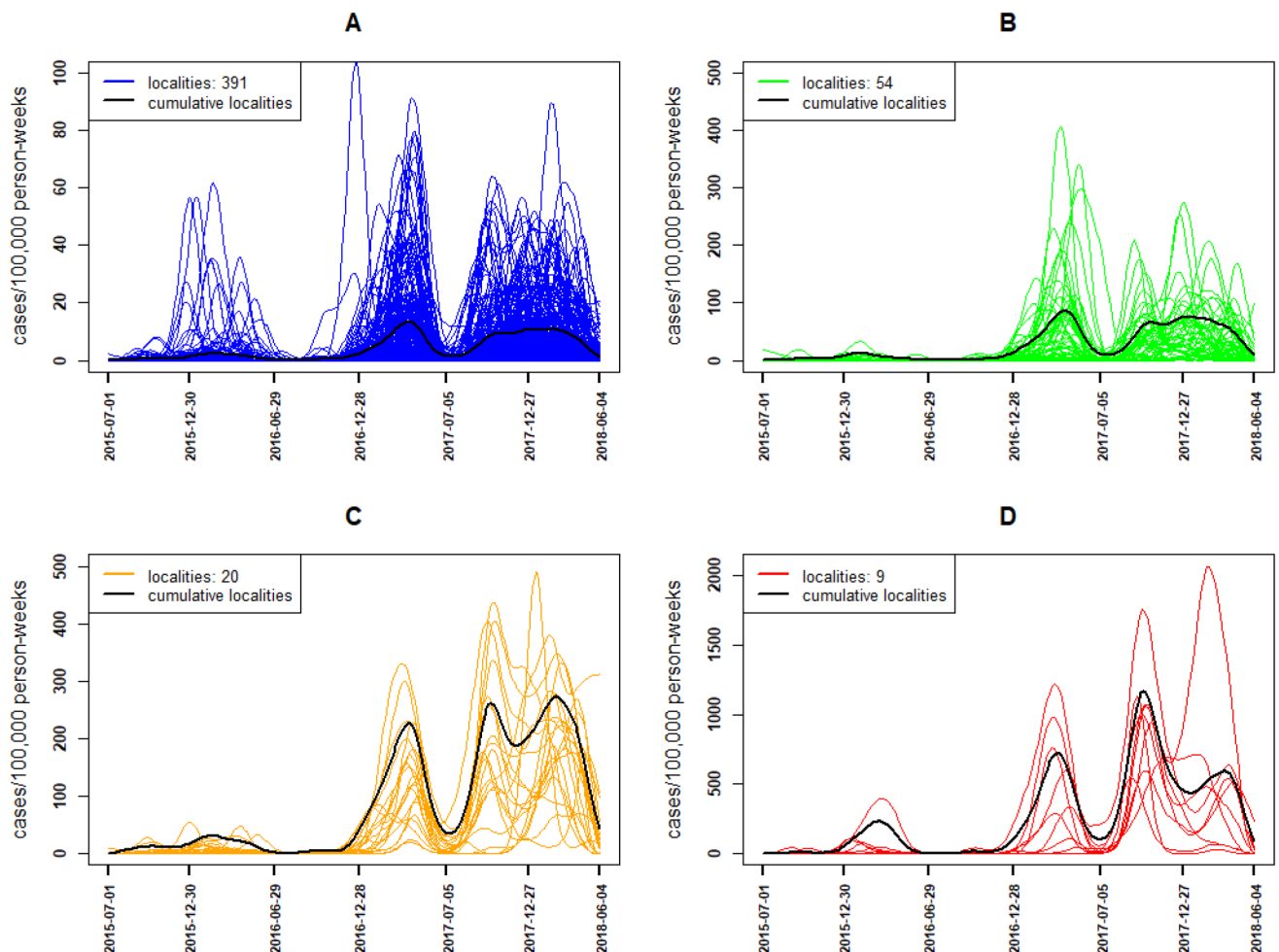


Figure 3. Smoothed incidence localities and cumulative incidence for each pattern from July 2015 to June 2018: Very low (A—blue), Low (B—green), Intermediate (C—orange) and High (D—red).

Spatial distribution of malaria pattern localities in Vhembe District, July 2015–June 2018

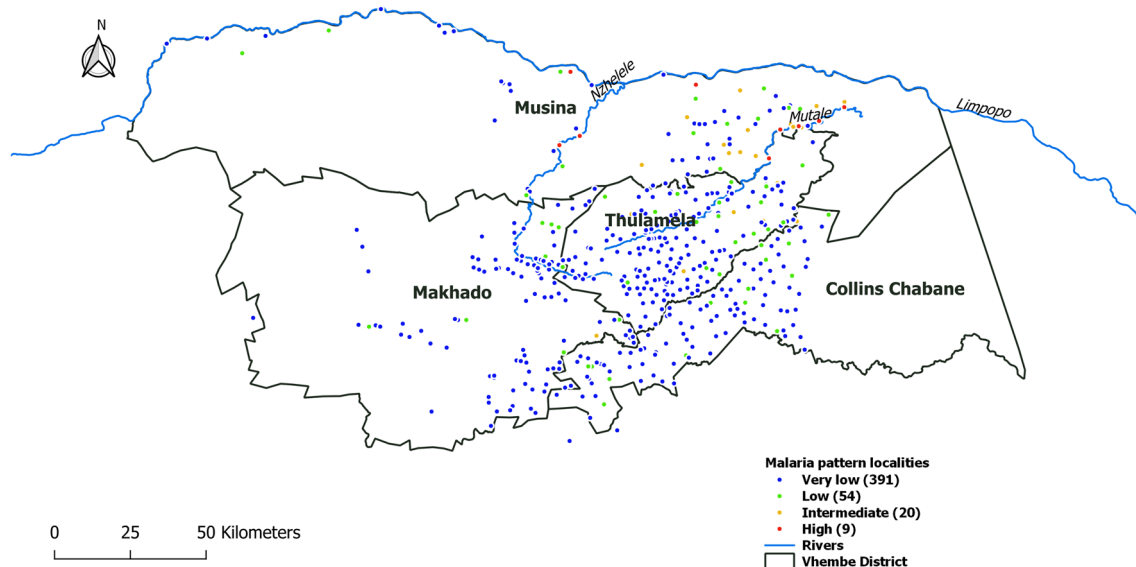


Figure 4. Spatial distribution of malaria incidence pattern localities in Vhembe District: very low (blue), low (green), intermediate (orange) and high (red). Maps were produced by the authors using QGIS software (version 3.10.1., QGIS Development Team 2020. Open-Source Geospatial Foundation Project) and Shapefiles downloaded from websites^{44,45}.

Characteristic ^a	Very low, N = 391	Low, N = 54	Intermediate, N = 20	High, N = 9
Locality population	1602 (640, 3300)	674 (209, 1876)	467 (258, 1290)	120 (100, 338)
Cumulative incidence	2.4	16.2	54.1	181.2
Locality incidence	3.0 (1.1, 6.2)	21.7 (14.7, 33.3)	57.3 (40.6, 71.0)	211.3 (144.8, 237.2)
Cumulative incidence during season 2015/2016	0.63	3.12	10.63	58.34
Cumulative incidence during season 2016/2017	3	16	44	144
Cumulative incidence during season 2017/2018	4	29	108	337
Locality incidence during season 2015/2016	0.0 (0.0, 0.8)	0.5 (0.0, 3.4)	10.2 (5.0, 18.1)	33.5 (0.0, 110.5)
Locality incidence during season 2016/2017	2 (1, 6)	21 (11, 33)	41 (36, 51)	125 (60, 217)
Locality incidence during season 2017/2018	5 (1, 10)	39 (27, 54)	111 (78, 143)	317 (294, 507)

Table 2. Characteristics of the four malaria incidence patterns. ^aIncidence in cases/1000 person-years; median (interquartile range).

Their seasonal peaks were more important during the beginning of the third malaria season 2017/2018 and their epidemic period were wider during this season.

Generally, very low and low patterns had the same temporal dynamics but at different levels of intensity, while intermediate and high patterns had slight differences at the end of the third season (Fig. 3). Furthermore, a general spatial trend was observed from south to north, with an increasing number of localities of the highest incidence patterns (from very low to high pattern) (Fig. 4).

Local resurgence in malaria incidence patterns. From July 2017 onwards (Fig. 3) some characteristics appear to be more accentuated in the high and intermediate patterns, such as the presence of local rebound or recrudescence. These two patterns each have two recrudescence onset points (local minimum) around July and December 2017 during the last season. In addition, malaria incidence quantiles increased across patterns (from very low to high pattern) and transmission seasons (from season 2015/2016 to 2017/2018).

Discussion

This work identified and described four different temporal malaria incidence patterns at the locality level in the Vhembe District during the three seasons; it further highlighted a local resurgence and an unexpected increase of malaria incidence during the study period.

In malaria endemic countries, as the elimination phase approaches, it is important to work on local spatial scales (villages and health posts, to name a few) and thin temporal scales (weekly)⁴⁶. Working with these fine scales make it possible to reconstruct the reality of the epidemic beyond the administrative limits over time. In addition, this will make it possible to avoid missing certain malaria particularities that could not be analysed on larger spatial and temporal scales. This is due to the heterogeneity between localities belonging to the same region and their differences on temporal dynamics, which is not helpful for refining targeted control actions in space and time for achieving malaria elimination⁴⁷.

Having data on these thin scales may be difficult in some countries or areas. However, in order to have relevant information for the development of control strategies, it would be necessary to conduct detailed analyses. This requires the implementation and effective monitoring of a data collection system^{12,48}. According to the WHO, there is no single intervention or set of interventions that will achieve malaria elimination in all countries. Rather, an arsenal of interventions tailored to the intensity and dynamics of malaria transmission in each country must be identified and used to achieve and sustain elimination. The effectiveness of interventions varies depending on where and when they are applied. In addition, accurate stratification of malaria intensity in local context is essential for effective targeting of interventions.

Therefore, the introduction of the functional approach in this framework of epidemiology may be an alternative. Indeed this approach answers simultaneously to a set of technical limits of classical methods^{49,50} as mainly the possibility to model simultaneously a large dataset with high dimension, the possibility to compare the spatio-temporal dynamic of malaria between them while analysing them in the same scale, the possibility to handle missing data or to deal with irregular observation times by working with estimated malaria incidence in real time during the study period, the possibility to appropriately classify malaria time series for clustering geographical units with the same intensities and same temporal dynamics^{16,17,51,52}. Indeed, it is important to capture how malaria incidence evolves over time across localities. Even if two localities have the same level of malaria incidence intensity, they would need different actions due to how they evolve over time. Moreover, the mathematical properties underlying this functional method also allow the analysis of malaria speed and acceleration corresponding to the first and second derivatives of the incidence curves. These derivatives make it possible to determine epidemiological indicators such as the dates of onset, acceleration, peak, slowdown and end of the epidemic^{16,53,54}. Although the latter points are not investigated in this study, it was important to show the potential of this method for further analysis.

One of the limitations of this approach, for operational aspects, could be the dispersion of the localities of the same pattern over the geographical space. However, this limitation can be overcome with good local coordination, i.e., at district or communal level, intervention strategies deployed on local targets and well organised priorities⁴⁶. Depending on the epidemiological realities of the countries, some characteristics would be different like the number of patterns, the temporal and spatial dynamic of patterns and epidemiological indicators^{16,17}.

Malaria transmission in South Africa is seasonal, with the low transmission period (winter) starting in June and ending in August; and the high transmission period (summer) starting in September and ending in May of the following year⁵. Generally, malaria cases start to increase in October, peak in January and February, and then decrease around May³¹. Our results showed that the last period of low transmission (June–August 2017), which was expected to have a very low level of incidence, had a sharp increase especially for the intermediate and high patterns and an unexpected peak around October 2017 in addition to the seasonal peak around February 2018: resurgence occurrence. It would be enlightening in this context to understand the factors that would be associated with the high intensity of incidence during the low transmission period of the 2017/2018 season and the occurrence of the resurgent series. These results open perspectives on the search for potential factors associated with these phenomena. However, several studies^{55,56} in South Africa has linked the recent malaria resurgence to climatic variables such as rainfall and temperature, to the mosquito abundance and arrival of new mosquito species. Under suitable conditions, mosquitoes bite more aggressively to produce more eggs, and in the process, transmit malaria faster than usual.

Furthermore, the spatial distribution of localities in Vhembe District showed that the further north one goes towards the Limpopo River (Mutale), the more one encounters malaria patterns with a high level of incidence indicating a spatial trend in malaria incidence. Specifically, the high pattern localities which are along the Mutale, Limpopo and Nzhelele Rivers (Fig. 4). One of them, Gwawkwani, which was an unrecognised hotspot, is currently benefiting from a research project to better understand the local risk factors⁵⁷. It has also been established that in South Africa low-lying areas (around 1000 m) are more affected by malaria^{5,58}, as is the case in the Limpopo River valley and its tributaries. In addition, this northern area borders on the malaria-endemic Zimbabwe and Mozambique⁵⁹, and imported cases are frequently reported due to economic migration during the rainy season (agriculture)²⁹. In terms of malaria control, South Africa has continued to apply indoor residual spraying as their primary means of control in the country¹⁰.

Studies have shown the relationship between climatic, environmental, entomological, preventive and social factors on malaria incidence in South Africa^{6,8,14}. The use of these data in this work could help to identify the factors associated with the locality malaria patterns and the dynamic characteristics of the factors associated with temporal increases and decreases (variations) in incidence in these localities. This latter point can be investigated by using functional regression^{18–20,40} that allows finer results on the effects of factors on the dynamic of malaria.

This study also proposed methodological tools for guiding intervention strategies by targeting priority areas and relevant dates for implementation and highlighted the need of:

- Improving the data collected at local scale across health information systems,
- Improving the availability of data population at local scale by official sources,

- Improving local environmental, economic and social surveillance for better investigating their effect at local scale,
- Additionally, defining and adapting health policies at local scale by:
 - Targeting priority areas using malaria patterns,
 - Identifying their dynamic characteristics and the factor effects over time,
- Implementing strategies according to their characteristics and identifying the appropriate times for each malaria pattern to apply the targeted control actions (allowed by the regularised smooth functions that estimated the incidence in real time even where they were not observed).

Conclusion

This study showed four temporal dynamics of malaria in Vhembe District with different characteristics. It also highlighted unusual malaria phenomena in Vhembe District such as a high level of incidence during a low transmission season and the occurrence of a series of resurgence. South Africa has reached the malaria pre-elimination stage, based on the development of access to diagnostics, treatment and vector control. However, the recent resurgence of malaria in remote areas hinders malaria elimination, highlighting the need for assessing the resurgence determinants and developing innovative targeted malaria elimination strategies in keeping with the local and global goals on eliminating malaria.

Data availability

The datasets analysed during the current study are available from the corresponding author on reasonable request.

Received: 18 September 2022; Accepted: 4 July 2023

Published online: 08 July 2023

References

1. World Health Organization. *World malaria report 2022* <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022> (2022).
2. World Health Organization. OMS|Bureau régional de l'Afrique. *WHO* (2020) <http://www.who.int/about/regions/afro/fr/>.
3. World Health Organization. WHO|Country profiles. *WHO* (2020). <http://www.who.int/malaria/publications/country-profiles/2017/en/>.
4. NICD—Disease Index—Malaria (2020). <https://www.nicd.ac.za/diseases-a-z-index/malaria/>.
5. National Institute for Communicable Disease. National guidelines for prevention of Malaria (2020). https://www.nicd.ac.za/wp-content/uploads/2019/03/National-Guidelines-for-prevention-of-Malaria_updated-08012019-1.pdf.
6. Maharaj, R. *et al.* Epidemiology of malaria in South Africa: From control to elimination. *SAMJ South Afr. Med. J.* **103**, 779–783 (2013).
7. Balawanth, R. *et al.* Assessing Kwa-Zulu-Natal's progress towards malaria elimination and its readiness for sub-national verification. *Malar. J.* **18**, 89 (2019).
8. Adeola, A. M. *et al.* Climatic variables and malaria morbidity in mutale local municipality, south africa: A 19-year data analysis. *Int. J. Environ. Res. Public Health* **14**, 89 (2017).
9. Hlongwana, K. W. & Tsoka-Gwegweni, J. From malaria control to elimination in South Africa: The researchers' perspectives. *Afr. J. Prim. Health Care Fam. Med.* **8**, 1 (2016).
10. Maharaj, R. *et al.* The feasibility of malaria elimination in South Africa. *Malar. J.* **11**, 423 (2012).
11. World Health Organization, Global Malaria Programme, & World Health Organization. *A Framework for Malaria Elimination* <https://www.who.int/publications/i/item/9789241511988> (2017).
12. Cox, J. *et al.* Novel approaches to risk stratification to support malaria elimination: An example from Cambodia. *Malar. J.* **13**, 371 (2014).
13. Abiodun, G. J. *et al.* A Dynamical and zero-inflated negative binomial regression modelling of malaria incidence in Limpopo Province, South Africa. *Int. J. Environ. Res. Public Health* **16**, 2000 (2019).
14. Adeola, A. *et al.* Rainfall trends and malaria occurrences in Limpopo Province, South Africa. *Int. J. Environ. Res. Public Health* **16**, 5156 (2019).
15. Cissoko, M. *et al.* Stratification at the health district level for targeting malaria control interventions in Mali. *Sci. Rep.* **12**, 8271 (2022).
16. Dieng, S. *et al.* Application of functional data analysis to identify patterns of malaria incidence, to guide targeted control strategies. *Int. J. Environ. Res. Public Health* **17**, 4168 (2020).
17. Legendre, E. *et al.* Malaria temporal dynamic clustering for surveillance and intervention planning. *Epidemics* **43**, 100682 (2023).
18. Ramsay, J. O. & Silverman, B. W. *Functional Data Analysis* (Springer, 2005).
19. Ramsay, J. O., Hooker, G. & Graves, S. *Functional Data Analysis with R and MATLAB* (Springer Science & Business Media, 2009).
20. Ramsay, J. O., Silverman, B. W., Ramsay, J. O. & Silverman, B. W. *Applied Functional Data Analysis: Methods and Case Studies* (Springer, 2002).
21. Ullah, S. & Finch, C. F. Applications of functional data analysis: A systematic review. *BMC Med. Res. Methodol.* **13**, 43 (2013).
22. Fleming, D. M., Zambon, M., Bartelds, A. I. M. & de Jong, J. C. The duration and magnitude of influenza epidemics: A study of surveillance data from sentinel general practices in England. *Wales Netherl.* **7**, 896. <https://doi.org/10.1023/a:1007525402861> (1999).
23. Rakocevic, B. *et al.* Influenza surveillance: Determining the epidemic threshold for influenza by using the Moving Epidemic Method (MEM), Montenegro, 2010/11 to 2017/18 influenza seasons. *Euro Surveill.* **24**, 8963 (2019).
24. Teklehaimanot, H. D., Schwartz, J., Teklehaimanot, A. & Lipsitch, M. Alert Threshold algorithms and malaria epidemic detection. *Emerg. Infect. Dis.* **10**, 1220–1226 (2004).
25. Vega, T. *et al.* Influenza surveillance in Europe: Comparing intensity levels calculated using the moving epidemic method. *Influenza Other Respir. Viruses* **9**, 234–246 (2015).
26. Outils DHIS2 pour la collecte et la compilation de données (2023). <https://www.who.int/fr/teams/global-malaria-programme/prevention/vector-control/dhis-data-collection-and-collation-tools>.
27. Jacques, J. & Preda, C. Functional data clustering: A survey. *Adv. Data Anal. Classif.* **8**, 231–255 (2014).
28. South Africa|Malaria Elimination 8 (2022). <https://malariaelimination8.org/south-africa>.

29. National Institute for Communicable Disease. Malaria elimination strategic plan for South Africa 2019–2023. *NICD* (2022). <https://www.nicd.ac.za/diseases-a-z-index/malaria/>.
30. Maharaj, R. *et al.* Decadal epidemiology of malaria in KwaZulu-Natal, a province in South Africa targeting elimination. *Malar. J.* **18**, 368 (2019).
31. Department of Health Limpopo province (2023). <http://www.doh.limpopo.gov.za/>.
32. Statistics South Africa, S. S. Statistics by place|Statistics South Africa (2021). http://www.statssa.gov.za/?page_id=964.
33. Frith, A. Census (2011). <https://census2011.adrianfrith.com/>.
34. Pelletier, F. & Spoorenberg, T. Séance 2 Aperçu sur les méthodes de projection, vol 36 https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/unpd_ws_201611_2_methodes_de_projection_final.pdf (2016).
35. Vhembe District Municipality (2022). http://za.geoview.info/limpopo/vhembe_district_municipality.
36. Murtagh, F. & Legendre, P. Ward's hierarchical agglomerative clustering method: Which algorithms implement ward's criterion?. *J. Classif.* **31**, 274–295 (2014).
37. Ward, J. H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–244 (1963).
38. Chouakria, A. D. & Nagabhushan, P. N. Adaptive dissimilarity index for measuring time series proximity. *ADAC* **1**, 5–21 (2007).
39. Montero, P. & Vilar, J. A. Tsclust: An r package for time series clustering. *J. Stat. Softw.* **62**, 1 (2014).
40. Febrero-Bande, M. *et al.* Statistical computing in functional data analysis: The R package fda. *usc. J. Stat. Softw.* **51**, 1–28 (2012).
41. Malouche, D. *Méthodes de Classifications* https://lmbp.uca.fr/DoWellB/docs/malouche/methodes_classifications_CF_Juin2013.pdf (2013).
42. Husson, F., Lê, S. & Pagès, J. *Exploratory Multivariate Analysis by Example Using R* (CRC Press, 2011).
43. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
44. GADM. *GADM Maps and Data* (2022). <https://gadm.org/>.
45. RQIS (Resource Quality Information Services)—Department of Water and Sanitation—South Africa—Home Page—formerly RQS(Resource Quality Services), formerly IWQS (Institute for Water Quality Studies), formerly HRI (Hydrological Research Institute) (2021). <http://www.dwa.gov.za/iwqs/>.
46. Gosling, R. *et al.* District-level approach for tailoring and targeting interventions: A new path for malaria control and elimination. *Malar. J.* **19**, 125 (2020).
47. Bejon, P. *et al.* Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya. *PLoS Med.* **7**, 1003 (2010).
48. El-Moamly, A. Malaria elimination: Needs assessment and priorities for the future. *J. Infect. Dev. Countr.* **7**, 769–780 (2013).
49. Helfenstein, U. Box-Jenkins modelling in medical research. *Stat. Methods Med. Res.* **5**, 3–22 (1996).
50. Kulldorff, M. A spatial scan statistic. *Commun. Stat. Theory Methods* **26**, 1481–1496 (1997).
51. Delsol, L. *Régression Sur Variable Fonctionnelle: Estimation, Tests de Structure et Applications* (Université Paul Sabatier-Toulouse III, 2008).
52. Ferraty, F. & Vieu, P. Richesse et complexité des données fonctionnelles. *Rev. Mod.* **43**, 25–43 (2011).
53. Guichard, D. 5. Curve Sketching (2022). https://www.whitman.edu/mathematics/calculus_online/chapter05.html.
54. Sunil Kumar Singh. Acceleration and deceleration—Kinematics fundamentals—OpenStax CNX (2010). <http://cnx.org/contents/f25d0bfc-5f61-411b-bcee-be8187ad5cc7@10>.
55. Abiodun, G. J. *et al.* Investigating the resurgence of malaria prevalence in South Africa Between 2015 and 2018: A scoping review. *Open Public Health J.* **13**, 129 (2020).
56. Ikeda, T. *et al.* Seasonally lagged effects of climatic factors on malaria incidence in South Africa. *Sci. Rep.* **7**, 2458 (2017).
57. Mabeba Moremi|Article|University of Pretoria (2022). <https://www.up.ac.za/up-one-health-for-change/article/2933814/mabeba-moremi>.
58. Gerritsen, A. A. M., Kruger, P., van der Loeff, M. F. S. & Grobusch, M. P. Malaria incidence in Limpopo Province, South Africa, 1998–2007. *Malar. J.* **7**, 162 (2008).
59. Gunda, R., Chimbari, M. J., Shamu, S., Sartorius, B. & Mukaratirwa, S. Malaria incidence trends and their association with climatic variables in rural Gwanda, Zimbabwe, 2005–2015. *Malar. J.* **16**, 393 (2017).

Acknowledgements

We thank the public health network “Réseau doctoral en santé publique” coordinated by the EHESP (School for Higher Studies in Public Health) for supporting the thesis project of SD. We thank the NGO PROSPECTIVE & COOPERATION for collaboration.

Author contributions

S.D. and J.G. designed the study, performed data processing, the statistical analysis and interpretation, and wrote the first draft of the article; T.C.A., T.K., M.R., S.P. and R.B. contributed the data collection and interpretations; S.T. and J.G. performed the data management and quality control; C.D.J., M.K.B. and H.T. contributed to the interpretation; R.B. contributed to the data collection, to the design study and to the interpretation of the results. All authors read and approved the final manuscript.

Funding

The project leading to this publication has received funding from the A*MIDEX Foundation of Aix-Marseille University, funded by socio-economic partners. The public health network “Réseau doctoral en santé publique” coordinated by the EHESP (School for Higher Studies in Public Health) for supporting the thesis project of SD, including her scientific internship at the UP- ISMC, University of Pretoria, Pretoria, South Africa.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023