RESEARCH
Open Access

CrossMark

Household level spatio-temporal analysis of *Plasmodium falciparum* and *Plasmodium vivax* malaria in Ethiopia

Dinberu Seyoum[1,2], Delenasaw Yewhalaw[3,4], Luc Duchateau[5], Patrick Brandt[6], Angel Rosas-Aguirre[1,7*†] and Niko Speybroeck[1*†]

Abstract

Background: The global decline of malaria burden and goals for elimination has led to an increased interest in the fine-scale epidemiology of malaria. Micro-geographic heterogeneity of malaria infection could have implications for designing targeted small-area interventions.

Methods: Two-year longitudinal cohort study data were used to explore the spatial and spatio-temporal distribution of malaria episodes in 2040 children aged < 10 years in 16 villages near the Gilgel-Gibe hydropower dam in Southwest Ethiopia. All selected households (HHs) were geo-referenced, and children were followed up through weekly house-to-house visits for two consecutive years to identify febrile episodes of *P. falciparum* and *P. vivax* infections. After confirming the spatial dependence of malaria episodes with Ripley's K function, SatScan[TM] was used to identify purely spatial and space-time clusters (hotspots) of annual malaria incidence for 2 years follow-up: year 1 (July 2008-June 2009) and year 2 (July 2009-June 2010).

Results: In total, 685 *P. falciparum* episodes (in 492 HHs) and 385 *P. vivax* episodes (in 290 HHs) were identified, representing respectively incidence rates of 14.6 (95% CI: 13.4–15.6) and 8.2 (95% CI: 7.3–9.1) per 1000 child-months at risk. In year 1, the most likely (128 HHs with 63 episodes, RR = 2.1) and secondary (15 HHs with 12 episodes, RR = 5.31) clusters of *P. vivax* incidence were found respectively in southern and north-western villages; while in year 2, the most likely cluster was located only in north-western villages (85 HHs with 16 episodes, RR = 4.4). Instead, most likely spatial clusters of *P. falciparum* incidence were consistently located in villages south of the dam in both years: year 1 (167 HHs with 81 episodes, RR = 1.8) and year 2 (133 HHs with 67 episodes, RR = 2.2). Space-time clusters in southern villages for *P. vivax* were found in August-November 2008 in year 1 and between November 2009 and February 2010 in year 2; while for *P. falciparum*, they were found in September-November 2008 in year 1 and October-November 2009 in year 2.

Conclusion: Hotspots of *P. falciparum* incidence in children were more stable at the geographical level and over time compared to those of *P. vivax* incidence during the study period.

Keywords: Spatio-temporal analysis, SatScan, *Plasmodium vivax*, *Plasmodium falciparum*, Active case detection, Ethiopia

* Correspondence: angelrosasa@gmail.com; niko.speybroeck@uclouvain.be
†Equal contributors
[1]Institute of Health and Society (IRSS), Université catholique de Louvain, Brussels, Belgium
Full list of author information is available at the end of the article

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 2 of 11

Background

Despite a decline in the global malaria burden over the past 15 years, about 3.5 billion people were at risk worldwide in 2015, and millions of them are still not accessing the services they need to prevent and treat malaria k. Of 438,000 registered malaria deaths in 2015, approximately 80% of the deaths were concentrated in just 15 countries, mainly in Africa [1]. According to the Ethiopian Ministry of Health [2], 2,174,707 malaria clinical cases and 662 deaths due to malaria were registered between September 2014 and August 2015 (Ethiopian fiscal year - EFY 2014/2015). Laboratory confirmation of malaria by either light microscopy (LM) or rapid diagnostic tests (RDTs) was performed in 1,867,059 (85.9%) clinical cases, showing a predominance of *Plasmodium falciparum* (63.7%) over *P. vivax* cases (36.3%). Oromia is the second regional state of Ethiopia with the highest malaria incidence, accounting for about 20% (430,969 cases) of total reported clinical cases in the country (430,969 cases), but the first one in terms of malaria mortality, representing about one third (214 deaths) of total malaria-related deaths in Ethiopia [2].

Malaria transmission is mainly seasonal and unstable throughout the country and varies due to differences in altitude, season, and population movement [3, 4]. A good understanding of the local epidemiology and transmission dynamics of malaria infections is key for better targeting the control measures [5–8]. Many factors have been reported to significantly influence malaria transmission in Ethiopia with likely different levels of interaction across space and time [9, 10]. Ecological factors facilitating breeding sites of *Anopheles arabiensis* (i.e. dams, irrigation canals, floods on shorelines, agricultural field puddles, wet lands, man-made pools, and rain pools) [11, 12] or resting places for adult mosquitos (i.e. surrounding vegetation, housing characteristics) are thought to be the main factors for malaria transmission [13]. Conditions that increase exposure to infectious mosquitos' bites (e.g. agriculture and livestock economic activities) [14], and human behavioural factors that limit the coverage and effectiveness of malaria control interventions (e.g. outdoor sleeping habits, low utilization of long-lasting insecticidal nets, poor treatment seeking behaviours, and low treatment adherence) may also influence the malaria risk [15].

A previous analysis of malaria surveillance data based on passive case detection (PCD) in villages located around the Gilgel-Gibe hydroelectric power dam in Southwest Ethiopia suggests different spatial and temporal variations of malaria episodes for both *P. falciparum* and *P. vivax* [10]. Until now the use of spatial-temporal tools to detect malaria hotspots (i.e. single villages or groups of households within villages with increased risk of malaria transmission) has not been applied to analyse malaria transmission in this area. Capitalising on the availability of two-year longitudinal malaria cohort data, this study explored the spatial and spatio-temporal distribution of *P. falciparum* and *P. vivax* malaria episodes in 2040 children aged < 10 years living in 16 villages around the Gilgel-Gibe hydropower dam. The study was conducted as part of several other studies, intended to assess the impact of the Gilgel-Gibe hydroelectric dam on health and other sectors (environment, agriculture and economy) following its starting operation in 2004 [16].
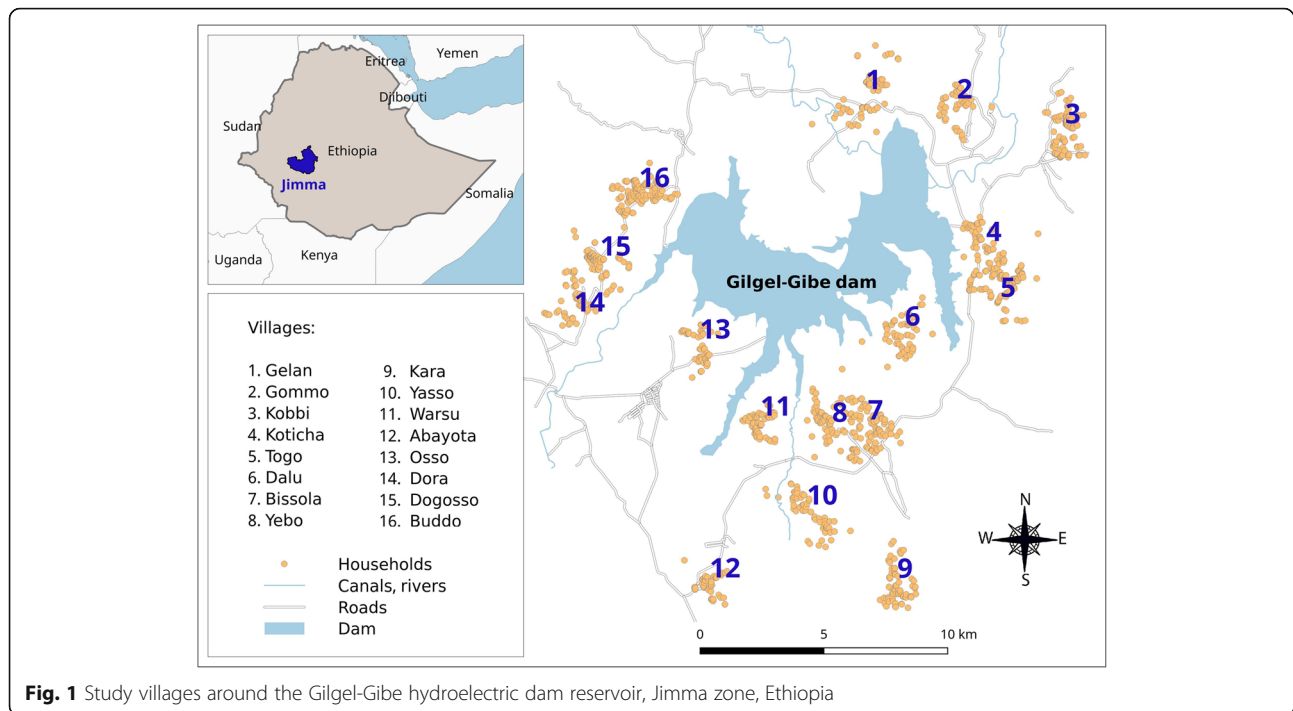
Methods

Study area

The study was conducted in Gilgel-Gibe dam area, in Jimma zone (Fig. 1), which is located 260 km south-west of Addis Ababa, in the Oromia region of Ethiopia. The study area lies between latitudes 7°42′50″N and 07°53′50″N and between longitudes 37°11′22″E and 37°20′36″E, at an altitude of 1734–1864 m above sea level. Sixteen villages within a 10 km radius (265–9046 m) from the dam reservoir shore were randomly selected based on similar eco-topography, access to health facilities, and homogeneity with respect to socio-cultural and economic activities [17, 18]. The main socio-economic activities of the households are mixed farming involving the cultivation of staple crops (maize, teff and sorghum), cattle, and small stock. All the households residing in the study villages belong to the Oromo ethnic group, which is one of the largest ethnic groups in Ethiopia [11].

Study design and population

A longitudinal 2-year malaria cohort study was conducted in children under 10 years old, living in the 16 selected villages around the Gilgel-Gibe hydroelectric power dam. A total of 2040 children aged less than 10 years were enrolled in July 2008 and completed weekly follow-ups until June 2010 [17]. Each child was identified with a unique code, and selected villages and households were geo-referenced using a handheld global positioning system (GPS) device (Garmin's GPSMAP 60CSx, Garmin International Inc., Olathe, Kansas, USA).

Follow-up, identification and management of malaria episodes

Active case detection (ACD) through weekly household visits was conducted to identify and register all febrile malaria episodes in the study population during the two-year follow-up period. During the household visits, axillary body temperature was taken, and the caregiver was asked about fever history. If a child had a fever (temperature ≥ 37.5 °C) or reported a history of fever in the past 24 h a finger-prick blood sample was taken for immediate diagnosis by LM in the same site, or at Omo-Nada District Health Center Laboratory. Trained

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 3 of 11

Fig. 1 Study villages around the Gilgel-Gibe hydroelectric dam reservoir, Jimma zone, Ethiopia

laboratory technicians conducted LM diagnosis. Thick smears were used to confirm the presence or absence of parasites, whereas the thin smear was used to identify the *Plasmodium* species.

All children with malaria confirmed by LM were treated according to the national treatment guidelines [19]. Treatment was administered by the parents and/or caregivers of the children, and consisted of 25 mg/kg of chloroquine (CQ) over three consecutive days for *P. vivax*, and artemether-lumefantrine (AL) for *P. falciparum* twice daily according to the body weight as follows: 5–14 kg, one tablet per dose; 15–24 kg, two tablets per dose; 25–34 kg, three tablets per dose; and adults, four tablets per dose. Treatment adherence was monitored during household visits by asking for the medication packages and the remaining pills. Absent children were followed in the next visits, and their caregivers were asked about the occurrence of symptomatic episodes and the confirmation of malaria at the health facility. In addition, all the health facilities near the study communities were visited monthly to verify their clinical records, checking if any enrolled children had presented a confirmed malaria episode in the past month, not being detected during the weekly visits.

Data analysis
Global spatial clustering
Data were double entered, validated and cleaned in Excel (Microsoft Corp, USA). A univariate Ripley's K-function was used to assess whether children with malaria episodes tended to be near other children with episodes

(determination of the expected number of children with episodes within distance h of a given child with episodes), while a bivariate K-function explored the spatial independence between children classified into two groups according to specific conditions (determination of the expected number of children with condition 1 within distance h of children with condition 2) [20].

For each malaria species and year of follow-up (first or second year), a file was created with the coordinates of children's households. The file was read in R software as a table, and a grid was created using the maximum and minimum values of latitude and longitude coordinates. After creating an object of class "ppp" for the point pattern distribution of children in a polygonal window, the Kest and Lest functions from the package Spatstat were applied to plot both observed and expected K-values over a range of distances [21]. Expected K-values and corresponding 95% confidence envelopes (CEs) were calculated using 999 Monte Carlo simulations to test the null hypothesis (H_o). The univariate K-function tested the H_o that the children with malaria episodes were randomly spatially distributed (complete spatial randomness, CSR), while the bivariate K-function tested two separated H_o: a) the children with and without malaria episodes were independently spatial distributed, and b) the children with malaria episodes, younger than 3 years old and older than 3 years old, were independently spatial distributed.

In the univariate analysis, observed K-values between low and high CE at specific distance h indicated random

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 4 of 11

distribution of the children with malaria episodes for that distance h, while those larger than the high CE values or those smaller than the low CE values at specific distance h indicated respectively significant spatial clustering or spatial dispersion of children with malaria episodes for that distance h. In the bivariate analysis, observed difference of K values between the children group 1 and group 2 which were between low and high CE at a specific distance h indicated spatial independence between the children groups for that distance h. Instead, a difference of K values larger than the high CE values at specific distance h indicated that the children group 1 tended to be more clustered than children group 2 for that distance h, while a difference smaller than the low CE values indicated that the household group 1 tends to be more dispersed than household group 2. In all analyses, distances h ranged from 0 to the maximum distance between the two closest children's households (about 5 km).

Local spatial clustering
The QGIS software v.2.16 (QGIS developer team, Open Source Geospatial Foundation) [22] was used to map all households with children aged less than 10 years old in the study area, and the SaTScan software v.9.3 (M Kulldorff and Information Management Services Inc, Boston, USA) [23, 24] was employed to detect spatial and space-time clusters of *P. falciparum* and *P. vivax* malaria episodes using the Bernoulli probability model [25]. The Bernoulli model in SaTScan requires input data as cases and controls. For each week of follow-up, cases were children with species-specific malaria episodes, while controls were children without episodes. Of note, malaria episodes were only counted for the week they were initially identified, and a child with a malaria episode (treated according to national guidelines) was censored for 21 days to prevent double counting of episodes in successive weeks.

The spatial analysis tests the null hypothesis of no clustering of children with malaria episodes. Different windows with varying size, from zero to a maximum radius of less than 15% of the total children, were allowed to move across the study area. This maximum radius was selected to avoid large non-populated areas within the identified malaria clusters by SaTScan. More details about the selection of maximum windows size can be found in Additional files 1 and 2. Each circle was a candidate cluster for which the log likelihood ratio (LLR) and the relative risk (RR) were obtained. The circular window with the highest LLR was defined as the most likely cluster (hotspot) if the P-value < 0.05 [24]. Once the hotspot was identified, a re-analysis of the children within that hotspot was conducted to identify whether that hotspot hid a smaller and more homogeneous area

with the highest malaria incidence (i.e. a hotspot within hotspot) [5].

The space-time analysis was performed under the null hypothesis that the risk of having malaria episodes was the same in all households and over time, with cylindrical windows having a circular geographic base and height corresponding to the time scale in weeks. The radius of each circular base was allowed to vary in size, to include up to as many as 15% of the total children. Comparably, the height of the cylinder varied in size up to a maximum of 50% of the study period with a time precision of one week. An unlimited number of overlapping cylinders with different dimensions were obtained, each cylinder corresponding to a possible space-time cluster. For each space-time cluster, the LLR was calculated and the most likely cluster defined as the cylinder with the maximum LLR. The statistical significance of the clusters was tested through 999 Monte Carlo simulations (the default value of the software) to achieve strong power, and the null hypothesis was rejected when the resulting p-value was below 0.05.

Results
Of the total 2040 followed-up children, 981 (48.1%) were females and 1059 (51.9%) males. The mean age at enrollment was 4.9 ± 2.0 years, not varying significantly across villages ($P > 0.05$) (Table 1). Of the total reported 1070 malaria episodes, 685 (363 episodes in year 1 and 322 episodes in year 2) were due to *P. falciparum* in 492 HHs, and 385 (296 episodes in year 1 and 89 episodes in year 2) were due to *P. vivax* in 290 HHs. *P. falciparum* incidence rates were respectively 15.5 and 13.7 episodes/ 1000 child-months in the first and second year of follow-up (14.6 episodes/1000 child-months thorough the study period), while *P. vivax* incidence rates were respectively 12.6 and 3.8 episodes/1000 child-months in the first and second year of follow-up (8.2 episodes/1000 child-months thorough the study period). Additional files 3 and 4 show videos with the spatio-temporal distribution of species-specific malaria incidence by HH. The visual inspection from first video suggests seasonal spatial distribution of *P. falciparum* incidence, with increased occurrence of *P. falciparum* episodes in households located at south of the dam mainly in the last months of the long rainy season (August and September), as well as, in the first months of the dry season (October and November). Instead, the second video shows a decreasing trend in the occurrence of *P. vivax* episodes along the study period, with a less clear spatial and seasonal pattern.

Global spatial clustering
Univariate K-function values for both species indicated that children with episodes were significantly clustered

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 5 of 11

Table 1 Baseline characteristics of enrolled children and incidence of malaria episodes by village

ID	Villages	Total	Age	Female	Number of episodes					
					P. falciparum			*P. vivax*		
		N	Mean ± SD	%	Year 1	Year 2	Total	Year 1	Year 2	Total
1	Gelan	133	4.9 ± 1.9	51.1	16	7	23	12	1	13
2	Gommo	127	5.3 ± 2.2	44.9	9	8	17	2	2	4
3	Kobbi	130	4.9 ± 1.7	43.8	5	7	12	2	2	4
4	Koticha	133	4.7 ± 1.8	51.1	16	9	25	11	1	12
5	Togo	124	5.2 ± 1.9	47.6	7	6	13	7	0	7
6	Dalu	136	4.5 ± 2.1	58.8	33	31	64	26	6	32
7	Bissola	134	4.8 ± 2.0	47.1	36	42	78	33	12	45
8	Yebo	126	5.3 ± 2.2	46.8	21	25	46	17	0	17
9	Kara	114	4.4 ± 1.9	42.1	38	32	70	27	8	35
10	Yasso	121	5.4 ± 2.1	43.8	26	29	55	30	11	41
11	Warsu	127	4.3 ± 2.1	43.3	27	16	43	27	8	35
12	Abayota	125	5.3 ± 2.1	56.0	34	30	64	20	4	24
13	Osso	129	5.6 ± 1.9	46.5	22	22	44	22	7	29
14	Dora	125	5.6 ± 1.4	42.4	17	11	28	16	3	19
15	Dogosso	127	5.3 ± 1.5	52.8	28	8	36	11	8	19
16	Buddo	129	4.0 ± 2.3	49.6	28	39	67	33	16	49
Total		2040	4.9 ± 2.0	48.1	363	322	685	296	89	385

Abbreviation: SD standard deviation

at all distances up to 5.0 km in both years of follow-up (Additional file 5). According to the bivariate K-function plots, the distribution of children with *P. falciparum* episodes in the first year were significantly more clustered than children without episodes only at distances greater than 4.0 km (Fig. 2a); while in the second year, this comparative increased clustering pattern of children with episodes occurred at all distances (Fig. 2b). Regarding *P. vivax*, children with episodes in the first year were significantly more clustered than those without episodes at distances lower than 1.5 km, and those larger than 3.5 km (Fig. 2c). In the second year, children with *P. vivax* episodes were significantly more clustered only at distances larger than 4.2 km (Fig. 2d). On the other hand, K-function analyses also indicated that children younger than 3 years and those older than 3 years, both with malaria episodes, were independently spatial distributed (Fig. 3).

Local spatial clustering

Purely spatial analysis by SaTScan confirmed that malaria episodes due to both species were not randomly distributed. The most likely spatial cluster of *P. falciparum* incidence in year 1 was a 5.5 km radius area south of the dam, composed of 167 HH presenting a total 81 episodes (Fig. 4a; Table 2). Households within this cluster belonged mainly to Kara and Yasso villages and were 1.8 times more at risk of acquiring *P. falciparum* infections

than households outside the cluster (RR = 1.8, *P* = 0.02). In year 2, the most likely cluster of *P. falciparum* episodes was also located south of the dam with a radius area of 3.1 km, including 133H of mainly Kara and Yasso villages and accounting for 67 episodes (RR = 2.2, *P* < 0.001) (Fig. 4b; Table 2). Interestingly, the re-analysis of the children within the most likely clusters for *P. falciparum* did not identify a further hotspot in both years. In addition, two secondary clusters were identified at south and west of the dam only in the second year (Fig. 4b; Additional file 6). Both the purely spatial (Fig. 4c) and the spatio-temporal analysis (Table 3) during the two-year period consistently confirmed the location of the most likely cluster of *P. falciparum* incidence south of the dam, with the latter analysis identifying 11 weeks with the highest incidence in year 1 (September 14th - November 29th 2008) and 5 weeks in year 2 (October 11th - November 14th 2009).

The most likely spatial clusters of *P. vivax* incidence were a 4.1 km radius area located south of the dam in year 1 (Fig. 4d; Table 2) and a 1.3 km radius area west of the dam in year 2 (Fig. 4e; Table 2). The first cluster included 128 HHs of mainly Yasso village and presenting 63 episodes (RR = 2.1, *P* = 0.004) while the second one included 88 HHs in Buddo village presenting 16 episodes (RR = 4.4, *P* = 0.008). The re-analysis of the children within the most likely clusters for *P. vivax* did not identify a further hotspot in both years. In addition, two
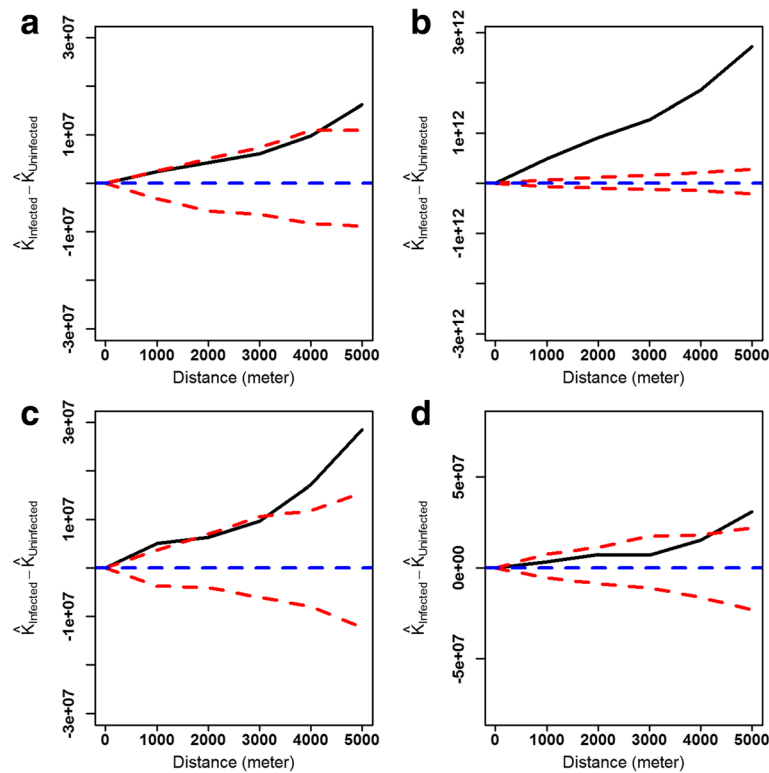
Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 6 of 11

Fig. 2 Bivariate Ripley's K function analysis comparing the spatial distribution of children with and without malaria episodes: **a** for *P. falciparum* incidence in the first year, **b** for *P. falciparum* incidence in the second year, **c** for *P. vivax* incidence in the first year, and **d** for *P. vivax* incidence in the second year. The *blue line* represents the expected difference of K function values ($K_{infected}$-$K_{noninfected}$) between children groups under the null hypothesis of spatial independence, the *solid black line* represents the observed difference K function, *dashed red lines* represents the confidence envelopes for expected K-function values calculated from 999 simulations
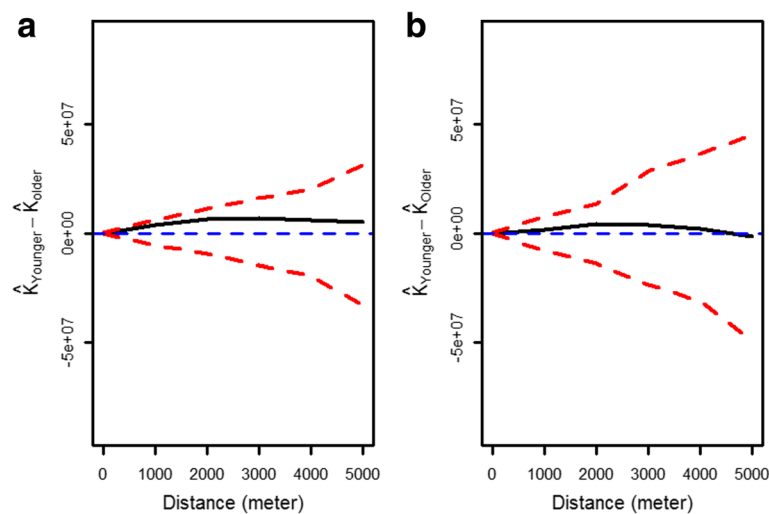

Fig. 3 Bivariate Ripley's K function analysis comparing the spatial distribution of two children groups who presented malaria cases, younger and older than 3 years old: **a** *P. falciparum* incidence during the two-year follow-up, **b** *P. vivax* incidence during the two-year follow-up. The *blue line* represents the expected difference of K function values ($K_{infected}$-$K_{noninfected}$) between children groups under the null hypothesis of spatial independence, the *solid black line* represents the observed difference K function, *dashed red lines* represents the confidence envelopes for expected K-function values calculated from 999 simulations
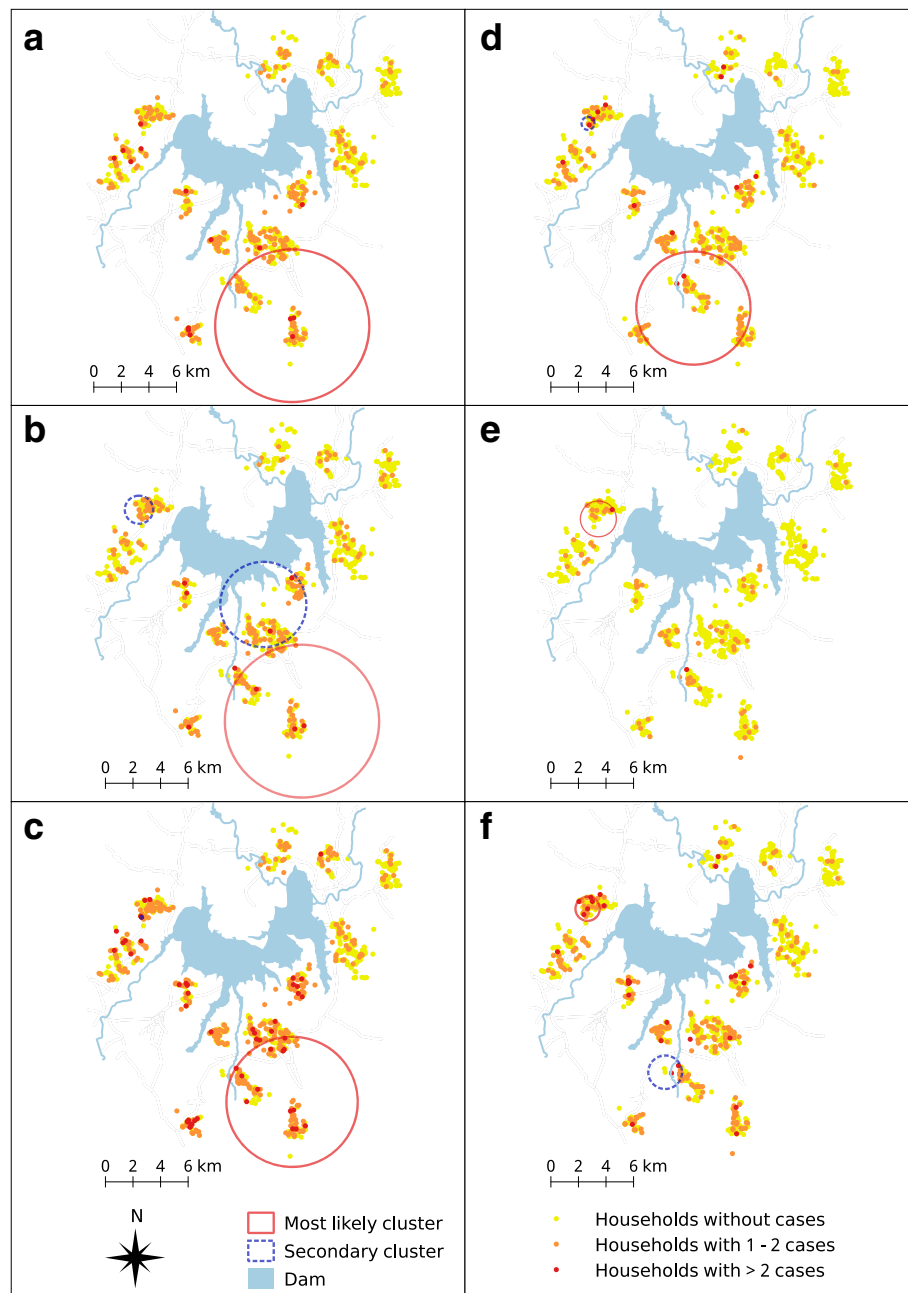
Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 7 of 11

Fig. 4 The most likely and secondary spatial clusters of *P. falciparum* and *P. vivax* incidence. *P. falciparum*: **a** (first year), **b** (second year) and **c** (both years); *P. vivax*: **d** (first year), **e** (second year), **f** (both years)

small secondary clusters were identified west of the dam only in year 1; one of them had a similar location to the most likely cluster identified in the following year (year 2) (Fig. 4d; Additional file 6). The spatio-temporal analysis confirmed the most likely clusters south and west of the dam respectively in year 1 and year 2. Fifteen weeks with highest *P. vivax* incidence were identified in year 1 (August 17th - November 29th 2008), and 13 weeks in year 2 (November 15th 2009 - February 13th 2010).

Discussion

Species-specific malaria episodes in children under 10 years old detected by ACD between July 2008 and June 2010 were clustered in groups of HHs (hotspots) in selected villages around the Gilgel-Gibe hydropower dam in Southwest Ethiopia. Like in other endemic Ethiopian regions [9, 26], *P. falciparum* episodes predominated over *P. vivax* episodes with *P. vivax* showing a sharp decrease in annual incidence rates during the study period.

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 8 of 11

Table 2 Spatial scan statistics of the most likely cluster of *P. falciparum* and *P. vivax* malaria episodes

	P. falciparum			*P. vivax*		
	Year 1	Year 2	Totals	Year 1	Year 2	Total
Coordinates (N, E)	7.6920 N, 37.3129E	7.6898 N, 37.3201E	7.7024 N, 37.3128E	7.7034 N, 37.2825E	7.8225 N, 37.2188E	7.8266 N, 37.2171E
Radius (km)	5.53	5.54	4.72	4.10	1.29	0.87
Households (%)	14.5% (167/1148)	11.6% (133/1148)	16.4% (189/1148)	11.1% (128/1148)	7.4% (85/1148)	6.01% (69/1148)
Population (%)	13.8% (282/2040)	10.5% (214/2040)	15.0% (306/2040)	11.6% (236/2040)	4.7% (96/2040)	3.8% (77/2040)
Cases	22.3% (81/363)	20.8% (67/322)	25.8% (177/685)	21.3% (63/296)	17.9% (16/89)	10.1% (39/385)
LLR	9.57	14.72	22.47	11.31	10.52	14.56
Relative risk	1.79	2.24	1.87	2.07	4.44	2.83
P-value	0.02	< 0.001	< 0.001	0.004	0.008	< 0.001

Abbreviation: *LLR* log likelihood ratio

Comparatively, hotspots of *P. falciparum* incidence in children were more stable at a geographical level and over time than those of *P. vivax* incidence, with consistent locations at the south of the dam in the two successive study years.

The level of statistical significance is an important factor in determining whether a certain geographical area forms a plausible hotspot of malaria transmission. In this study, the global clustering K-function test suggested the existence of clustering of children with malaria episodes without pinpointing specific locations, while its variant, the bivariate K-function, was able to demonstrate that children with malaria episodes tended to be more aggregated than children without episodes. Although a previous study using recurrent-event models to analyse incidence data in the same study children showed contrasting associations between the age of children and the species-specific malaria incidence (i.e. *P. vivax* were mostly observed in younger age groups, while *P. falciparum* episodes were mainly seen in older children) [18], the age would not have any influence over the spatial distribution of children presenting species-specific malaria episodes according to the bivariate K-function analysis.

In contrast to global clustering tests, local clustering tests (i.e. the Kulldorf spatial scan statistic) were able to identify the most likely location of hotspots of malaria incidence in the two consecutive years of study. Indeed, hotspots of *P. falciparum* incidence suggested a higher exposure to infectious mosquitoes in southern villages, especially after the long rainy season (peak of cases between September and November according to space-time analysis). After rains, intermittent streams could create pockets or pools of water which can serve as potential breeding sites for mosquitoes *Anopheles arabiensis*, contributing to an increase in mosquito density and vector-human contacts, and consequently to a greater number of malaria episodes during the dry season [18, 27]. The characteristics of the southern land (i.e. wet, flat and silted) [28], as well as, the landslides that often occur there [29, 30], would additionally increase the accumulation of water in shallow pits that act as excellent mosquito breeding habitats. Of note, as previously found in a recent article [18], the Gilgel-Gibe dam reservoir would not have a significant impact on the malaria transmission in the study area, since the design and automatic operation of the dam would be able to prevent the appearance

Table 3 Space-time scan statistics of the most likely cluster of *P. falciparum* and *P. vivax* malaria episodes

	P. falciparum		*P. vivax*	
	Year 1	Year 2	Year 1	Year 2
Time period (initial - final)	14/09/08–29/11/08	11/10/09–14/11/09	17/08/08–29/11/08	15/11/09–13/02/10
Coordinates (N, E)	7.6806N, 37.2524E	7.7391N, 37.3161E	7.7002N, 37.3172E	7.7243N, 37.2763E
Radius (km)	6.83	4.09	5.22	2.19
Households (%)	12.6% (145/1148)	16.2% (186/1148)	16.9% (194/1148)	4.9% (56/1148)
Population (%)	14.4% (295/2040)	14.5% (297/2040)	14.8% (302/2040)	5.11% (104/2040)
Cases	9.3% (34/363)	15.2% (49/322)	15.9% (47/296)	12.3% (11/89)
LLR	20.97	73.70	29.57	15.09
Relative risk	4.03	11.90	4.22	10.24
P-value	< 0.001	< 0.001	< 0.001	0.012

Abbreviation: *LLR* log likelihood ratio

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 9 of 11

of shoreline puddles and consequently the formation of mosquito breeding sites near the reservoir.

In contrast to *P. falciparum*, the hotspots of *P. vivax* incidence were less stable in place and time during the study period, suggesting that the occurrence of *P. vivax* clinical episodes are less sensitive to seasonal and environmental changes than *P. falciparum*, and that other factors should be considered to understand the spatial-temporal heterogeneity of infections due to this species [18]. The biological features of *P. vivax* infections may also influence the spatial distribution of the disease, particularly the ability of parasites to relapse weeks or months after a primary parasitaemia [31]. However, the characterization and prediction of spatial patterns remain challenging because of the difficulty distinguishing between a hypnozoite-triggered relapse, a resurgence of erythrocytic parasites (i.e. recrudescence) due to a failure in the therapy, or reinfection of an individual with a new parasite strain following a primary infection [32, 33]. This challenge is even greater considering that children with confirmed *P. vivax* episodes in the study received chloroquine (CQ) but not primaquine (PQ), following the national guidelines for areas where the glucose-6-phosphate dehydrogenase deficiency (G6PD deficiency) is not known, and where tests to detect that condition are not available [19].

The immunity to the malaria infection is another factor that should be considered when interpreting the spatial and spatial-time patterns of malaria clinical episodes in study children. As previously hypothesised in a recent article [18], differences in clinical malaria incidence between species with respect to age may be related to different species-specific acquisition rates of immunity [34], with immunity acquired more rapidly with *P. vivax* than with *P. falciparum*. Taking this into account, the fast development of clinical immunity in areas with the highest *P. vivax* exposure and incidence (i.e. southern villages) identified during the first study year may also explain why the hotspots did not remain in the same location in the following year. Similarly, immunity would be a factor to be considered in the analysis of the reduction of *P. vivax* clinical incidence rates during the study period [18].

The main limitations of our study may be related to the malaria metrics and the geospatial analysis used for the identification of clusters of malaria transmission in the area. The spatial analysis of clinical malaria incidence obtained through rigorous weekly active case detection of symptomatic episodes in enrolled children may be the best method for detecting malaria hotspots if most malaria infections occurred in those children were symptomatic and microscopically detected [6]. However, this cannot be confirmed in this study because the methodology did not include the screening for asymptomatic and sub-microscopic malaria infections. Malaria surveys in other endemic regions of Ethiopia have reported that asymptomatic and sub-microscopic infections can represent an important proportion of total malaria infections [35, 36]; however, the cross-sectional design in the latter studies did not take into account the incubation period for some of those infections, hence the potential development of symptoms and the increase of parasite density levels at a later stage [37], which would facilitate their detection by surveillance methods with strict follow-up like those included in our study. Further research is needed to better understand the impact of asymptomatic and sub-microscopic infections in endemic areas of Ethiopia, and to assess whether their spatial distribution differs from that of symptomatic infections. On the other hand, despite the recognition of SaTScan as a powerful tool to analyse spatial patterns of vector-borne diseases such as malaria [38, 39], a number of studies have pointed out that setting critical parameters in the analysis such as the maximum window size is not straightforward and suggested that this task should consider the application goals of the cluster detection and geographic scale of processes leading to the clusters [40]. Following these suggestions, our analysis set the maximum window size at 15%, instead of the default value of 50%, with the purpose of detecting the zones with the highest malaria incidence within the entire study area. This chosen parameter value did not only avoid large non-populated areas within circular clusters but also reduced (without eliminating) the influence of the uneven inter- and intra-village distribution of children/households in the study area. Moreover, the no identification of smaller and more homogeneous clusters with even higher transmission among the most likely malaria clusters of children/households (i.e. no more hotspots within hotspots), further supports our selection for the maximum window size, as well as the validity of the, detected most likely clusters for *P. falciparum* and *P. vivax* incidence (despite being composed of most households of a village or a group of contiguous villages). Other limitations of the study may be related to the absence of data on other potential risk factors for malaria infection (e.g. household size, education level in parents, malaria prevention practices at household, treatment-seeking behaviour, vegetation coverage, etc.) preventing the evaluation of their influence on the spatial distribution of the clinical malaria incidence. Gender and distance to the Gilgel-Gibe dam were not considered as covariates for the spatial analysis since those variables were not significantly associated with both *P. vivax* and *P. falciparum* malaria incidence in a longitudinal modelling approach [18]; while the children's age (despite being associated with malaria incidence according to the same model), which was found to be evenly distributed across villages, may not have influence over the spatial distribution of children

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 10 of 11

with malaria episodes in the study area according to the K-function test, hence not meeting the criteria to be catalogued as covariate [24].

Conclusion

Hotspots of *P. falciparum* incidence in children were more stable at the geographical level and over time than *P. vivax* incidence in selected villages around the Gilgel-Gibe hydropower dam in Southwest Ethiopia. Different malaria spatial-time patterns due to *Plasmodium falciparum* and *Plasmodium vivax* should be taken into account to better design and deliver targeted interventions.

Additional files

Additional file 1: Selection of window size for SaTScan analysis. (DOCX 17 kb)

Additional file 2: Spatial scan statistics of the most likely cluster of malaria episodes using different maximum window sizes. (DOCX 16 kb)

Additional file 3: Spatio-temporal distribution of households presenting *P. falciparum* episodes. Each yellow point represents an individual household. When a *P. falciparum* episode in a child is found in the household, the point becomes red. (MP4 862 kb)

Additional file 4: Spatio-temporal distribution of households presenting *P. vivax* episodes. Each yellow point represents an individual household. When a *P. vivax* episode is found in a child in the household, the point becomes red. (MP4 769 kb)

Additional file 5: Ripley's K function analyses showing significant spatial clustering of children with malaria episodes. *P. falciparum*: a (first year), b (second year); *P. vivax*: c (first year), d (second year). The red line represents the expected K function values under the null hypothesis of complete spatial randomness, the solid black line represents observed K function values, and the grey area represents confidence envelopes for expected K-function values calculated from 999 simulations. (TIF 7910 kb)

Additional file 6: Spatial scan statistics of the secondary clusters of *P. falciparum* and *P. vivax* malaria episodes by year of study. (DOCX 15 kb)

Abbreviations
ACD: Active case detection; AL: Artemether-lumefantrine; CE: Confidence envelopes; CI: Confidence interval; CQ: Chloroquine; CSR: Complete spatial randomness; EFY 2014/2015: Ethiopian fiscal year 2014/2015; GPS: Global positioning system; HH: Household; LLR: Log likelihood ratio; LM: Light microscopy; PCD: Passive case detection; RDT: Rapid diagnostic test; RR: Relative risk

Availability of data and material
Datasets are available from the corresponding author on reasonable request.

Authors' contributions
DS involved in the study design and performed the data cleaning, statistical analysis and drafted the manuscript; DY conceived and designed the study and reviewed the manuscript; LD was involved in designing the idea of the study and critically reviewed the manuscript; PB contributed to draft the manuscript, read critically and approved the final manuscript; AR contributed to design the study and the analysis, drafted the manuscript, read critically and approved the final manuscript; NS designed the study, was involved in the analysis, coordinated and critically reviewed and revised the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Ethical approval for the study was obtained from Jimma University Research and Ethics Committee. Verbal and written signed informed consent was obtained from the mother or caregiver of each child before enrollment of the child in the study.

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

Author details
[1]Institute of Health and Society (IRSS), Université catholique de Louvain, Brussels, Belgium. [2]Department of Statistics, Natural Science College, Jimma University, Jimma, Ethiopia. [3]Department of Medical Laboratory Sciences and Pathology, College of Health a Sciences, Jimma University, Jimma, Ethiopia. [4]Tropical and Infectious Diseases Research Center, Jimma University, Jimma, Ethiopia. [5]Department of Comparative Physiology and Biometrics, Faculty of Veterinary Medicine, Ghent University, Gent, Belgium. [6]School of Economic, Political and Policy Sciences, The University of Texas, Dallas, USA. [7]Institute of Tropical Medicine "Alexander von Humboldt", Universidad Peruana Cayetano Heredia, Lima, Peru.

References
1. World Health Organization. World malaria report 2015. Geneva: WHO; 2015.
2. Federal Democratic Republic of Ethiopia Ministry of Health. HSDP IV annual performance report 2014/2015. Addis Ababa: Ministry of Health; 2015.
3. Ghebreyesus TA, Haile M, Witten KH, Getachew A, Yohannes M, Lindsay SW, et al. Household risk factors for malaria among children in the Ethiopian highlands. Trans R Soc Trop Med Hyg. 2000;94:17–21.
4. Woyessa A, Deressa W, Ali A, Lindtjørn B. Prevalence of malaria infection in Butajira area, south-central Ethiopia. Malar J. 2012;11:84.
5. Bejon P, Williams TN, Nyundo C, Hay SI, Benz D, Gething PW, et al. A micro-epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within hotspots. Elife. 2014;3:e02130.
6. Rosas-Aguirre A, Speybroeck N, Llanos-Cuentas A, Rosanas-Urgell A, Carrasco-Escobar G, Rodriguez H, et al. Hotspots of malaria transmission in the Peruvian Amazon: rapid assessment through a parasitological and serological survey. PLoS One. 2015;10:e0137458.
7. Rosas-Aguirre A, Llanos-Cuentas A, Speybroeck N, Cook J, Contreras-Mancilla J, Soto V, et al. Assessing malaria transmission in a low endemicity area of north-western Peru. Malar J. 2013;12:339.
8. Alemu K, Worku A, Berhane Y. Malaria infection has spatial, temporal, and spatiotemporal heterogeneity in unstable malaria transmission areas in Northwest Ethiopia. PLoS One. 2013;8:e79966.
9. Sena L, Deressa W, Ali A. Dynamics of *Plasmodium falciparium* and *Plasmodium vivax* in a micro-ecological setting, southwest Ethiopia: effects of altitude and proximity to a dam. BMC Infect Dis. 2014;14:625.
10. Alemu K, Worku A, Berhane Y, Kumie A. Spatiotemporal clusters of malaria cases at village level, northwest Ethiopia. Malar J. 2014;13:223.
11. Kibret S, Wilson GG, Tekie H, Petros B. Increased malaria transmission around irrigation schemes in Ethiopia and the potential of canal water management for malaria vector control. Malar J. 2014;13:360.
12. Degefa T, Zeynudin A, Godesso A, Michael YH, Eba K, Zemene E, et al. Malaria incidence and assessment of entomological indices among resettled communities in Ethiopia: a longitudinal study. Malar J. 2015;14:24.
13. Dejenie T, Yohannes M, Assmelash T. Characterization of mosquito breeding sites in and in the vicinity of tigray microdams. Ethiop J Health Sci. 2011;21:57–66.

Seyoum *et al. Parasites & Vectors* (2017) 10:196

Page 11 of 11

14. Yimer F, Animut A, Erko B, Mamo H. Past five-year trend, current prevalence and household knowledge, attitude and practice of malaria in Abeshge, south-central Ethiopia. Malar J. 2015;14:230.

15. Sena LD, Deressa WA, Ali AA. Predictors of long-lasting insecticide-treated bed net ownership and utilization: evidence from community-based cross-sectional comparative study, Southwest Ethiopia. Malar J. 2013;12:406.

16. Yewhalaw D, Kassahun W, Woldemichael K, Tushune K, Sudaker M, Kaba D, et al. The influence of the Gilgel-Gibe hydroelectric dam in Ethiopia on caregivers' knowledge, perceptions and health-seeking behaviour towards childhood malaria. Malar J. 2010;9:47.

17. Yewhalaw D, Getachew Y, Tushune K, Michael K, Kassahun W, Duchateau L, et al. The effect of dams and seasons on malaria incidence and anopheles abundance in Ethiopia. BMC Infect Dis. 2013;13:161.

18. Seyoum D, Kifle GY, Rondeau V, Yewhalaw D, Duchateau L, Rosas-Aguirre A, et al. Identification of different malaria patterns due to *Plasmodium falciparum* and *Plasmodium vivax* in Ethiopian children: a prospective cohort study. Malar J. 2016;15:208.

19. Federal Democratic Republic of Ethiopia Ministry of Health. M of H. National malaria guideline. 3rd ed. Addis Ababa: Ministry of Health; 2012.

20. Gatrell AC, Bailey TC, Diggle PJ, Rowlingson BS. Spatial point pattern analysis and its application in geographical epidemiology. Trans Inst Br Geogr. 1996;21:256–74.

21. Baddeley A, Turner R. spatstat: an R package for analyzing spatial point patterns. J Stat Softw. 2005;12(6):1–42.

22. QGIS developer team. QGIS User Guide Release 2.14. Open Source Geospatial Foundation; 2014. http://docs.qgis.org/2.14/en/docs/user_manual/. Accessed 30 Nov 2016.

23. Kulldorff M. A spatial scan statistic. Commun Stat - Theory Methods. 1997;26:1481–96.

24. Kulldorff M. SaTScan user guide for version 9.0. Boston: M Kulldorff and Information Management Services Inc; 2011.

25. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. Stat Med. 1995;14:799–810.

26. Alemu A, Abebe G, Tsegaye W, Golassa L. Climatic variables and malaria transmission dynamics in Jimma town, South West Ethiopia. Parasit Vectors. 2011;4:30.

27. Abate A, Degarege A, Erko B. Community knowledge, attitude and practice about malaria in a low endemic setting of Shewa Robit Town, northeastern Ethiopia. BMC Public Health. 2013;13:312.

28. Devi R, Tesfahune E, Legesse W, Deboch B, Beyene A. Assessment of siltation and nutrient enrichment of Gilgel Gibe dam, Southwest Ethiopia. Bioresour Technol. 2008;99:975–9.

29. Broothaerts N, Kissi E, Poesen J, Van Rompaey A, Getahun K, Van Ranst E, et al. Spatial patterns, causes and consequences of landslides in the Gilgel Gibe catchment, SW Ethiopia. Catena. 2012;97:127–36.

30. Woldearegay K. Review of the occurrences and influencing factors of landslides in the highlands of Ethiopia: With implications for infrastructural development. Momona Ethiop J Sci. 2013;5:3–31.

31. White NJ. Determinants of relapse periodicity in *Plasmodium vivax* malaria. Malar J. 2011;10:297.

32. Mueller I, Galinski MR, Baird JK, Carlton JM, Kochar DK, Alonso PL, et al. Key gaps in the knowledge of *Plasmodium vivax*, a neglected human malaria parasite. Lancet Infect Dis. 2009;9:555–66.

33. Wells TNC, Burrows JN, Baird JK. Targeting the hypnozoite reservoir of *Plasmodium vivax*: the hidden obstacle to malaria elimination. Trends Parasitol. 2010;26:145–51.

34. Lin E, Kiniboro B, Gray L, Dobbie S, Robinson L, Laumaea A, et al. Differential patterns of infection and disease with P. falciparum and P. vivax in young Papua New Guinean children. PLoS One. 2010;5:e9047.

35. Santana-Morales MA, Afonso-Lehmann RN, Quispe MA, Reyes F, Berzosa P, Benito A, et al. Microscopy and molecular biology for the diagnosis and evaluation of malaria in a hospital in a rural area of Ethiopia. Malar J. 2012;11:199.

36. Golassa L, Enweji N, Erko B, Aseffa A, Swedberg G. Detection of a substantial number of sub-microscopic *Plasmodium falciparum* infections by polymerase chain reaction: a potential threat to malaria control and diagnosis in Ethiopia. Malar J. 2013;12:352.

37. Rosas-Aguirre A, Erhart A, Llanos-Cuentas A, Branch O, Berkvens D, Abatih E, et al. Modelling the potential of focal screening and treatment as elimination strategy for *Plasmodium falciparum* malaria in the Peruvian Amazon Region. Parasit Vectors. 2015;8:261.

38. Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, et al. Hitting hotspots: spatial targeting of malaria for control and elimination. PLoS Med. 2012;9:e1001165.

39. Bousema T, Drakeley C, Gesase S, Hashim R, Magesa S, Mosha F, et al. Identification of hot spots of malaria transmission for targeted malaria control. J Infect Dis. 2010;201:1764–74.

40. Chen J, Roth RE, Naito AT, Lengerich EJ, Maceachren AM. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. Int J Health Geogr. 2008;7:57.