ISSN: 2474-3658

Aserss et al. J Infect Dis Epidemiol 2024, 10:327
DOI: 10.23937/2474-3658/1510327
Volume 10 | Issue 9
Open Access

Journal of
Infectious Diseases and Epidemiology

RESEARCH ARTICLE

Development and Validation of an Empirical Model to Forecast Malaria Outbreaks at Amhara Region, Ethiopia: A Retrospective Follow-Up Study

Fetlework Workineh Aserss[1], Worku Awoke[1], Zelalem Mehari[1] and Damtie Lankir Abebe[2*]

[1]*Department of Epidemiology and Biostatistics, School of Public Health, College of Medicine and Health Sciences, Bahir Dar University, Bahir Dar, Ethiopia*

[2]*Amhara National Regional State Public Health Institute, Public Health Emergency Management Directorate, Bahirdar, Ethiopia*

***Corresponding author:** *Damtie Lankir Abebe, Amhara National Regional State Public Health Institute, Public Health Emergency Management Directorate, Bahirdar, Ethiopia*

Abstract

Background: Malaria remains a significant public health concern in developing countries. Distinct geographical regions have different factors that influence malaria transmission. The aim of this study is to develop and validate an empirical model to forecast malaria outbreak.

Method: A retrospective follow-up study was conducted from June 01 to 30, 2022 in thirty-four woredas that have metrological stations and nine developmental corridors. The collected data was analyzed by R version 4.0.4.Backward stepwise multivariable logistic regression was used.

Result: The presence of irrigation (OR = 1.522, 95% CI = 1.161-2.142), sunshine (SH > = 7.167) (OR = 4.104, 95% CI = 1.706-9.791), Rainfall (> = 98.178) (OR = 21.73, 95% CI = 5.755-141.326) and minimum temperate (OR = 0.956, 95% CI = 0.956-0.997) were significantly associated with malaria outbreaks.

Conclusions: Sunshine, minimum temperature, rainfall, and irrigation were important to forecasting malaria outbreak. Current month climate data have the fitted predictor to forecast the outbreak of malaria.

Keywords

Malarial outbreak, Forecast, Model

Abbreviations

AOR: Adjusted Odds Ratio; APHI: Amhara Public Health Institute; AUC: Area under the Curve; CI: Confidence Interval; MICE: Multiple Imputation with Chained Equations; SAF: Seasonal Adjusted Factors; PF: *Plasmodium falciparum*; PV: *Plasmodium vivax*; RDT: Rapid Diagnostic Test; SPSS: Statistical Package for the Social Science

Background

A sudden rise in the number of cases of a disease above what is normally expected in that population in that area is called a malaria outbreak [1-3]. Based on severity and progression depending on the type of pathogen causing the outbreak, the effectiveness of early warning systems, the preparedness and response time, and cultural and security factors has several stages. The stages of an outbreak include pre-outbreak, individual cases and small clusters of disease, widespread disease, outbreak control, and post-outbreak. If an outbreak of a highly infectious disease is suspected, this should be reported to the appropriate national and international authorities as stated by the International Health Regulations [4].

Climate-based distribution model of malaria transmission in Sub-Saharan Africa describes a simple numerical approach to defining the distribution of malaria transmission, based on biological constraints of climate on parasite and vector development. It provides a numerical basis for further refinement and prediction of the impact of climate change on transmission. Together with population, morbidity, and mortality data, the model provides a fundamental tool for the strategic control of malaria [5].

Malaria is key public health challenges in Ethiopia.

Citation: Aserss FW, Awoke W, Mehari Z, Abebe DL (2024) Development and Validation of an Empirical Model to Forecast Malaria Outbreaks at Amhara Region, Ethiopia: A Retrospective Follow-Up Study. J Infect Dis Epidemiol 10:327. doi.org/10.23937/2474-3658/1510327

Accepted: September 17, 2024: **Published:** September 19, 2024

It is one of the significant public health emergencies that cause a high-level of morbidity in Amhara region. The regional landmass is favorable to malaria transmission. Drivers of malaria transmission vary across different geographical regions. Climatic variables are major risk factor in seasonal and secular patterns of malaria transmission along Amhara; Ethiopia. There is widespread transmission of *Plasmodium falciparum (PF)* and *Plasmodium vivax (PV*) malaria, with a ratio of 1.2 of *PF* to *PV*, as seen in blood film tests from a cross-sectional survey [6].

An empirical model is a model where it is determined by the observed relationship among experimental data. In developing a correlation, it needs first to identify all the variables that may have an influence on it. However, in addition to their capacity for process optimization, empirical models can be utilized for calibration and experimental data prediction for a given system. Empirical models that have been used for the handling of weathering data have typically used curve fitting processes to generalize the results of experiment [7]. It is a useful tool for interpreting and applying surveillance data. It has a great potential to be used as a decision-support tool to predict mosquito-borne disease outbreaks [8,9]. The objective of this study is to develop and validate an empirical model to forecast malaria outbreak at Amhara region, 2022.

Methods

Study setting

The study was conducted in Amhara national regional state which has a metrological agency and has a five-year malaria surveillance data of 34 districts.

Study design and periods

A retrospective study design was conducted from June 1 to 30, 2022 in Amhara regional state, Northwest Ethiopia. The start and end of the recruitment period for this study covered January 1, 2016, to December 30, 2021.

Study population

The study population was all reported cases of malaria surveillance data in the Amhara public health institute within five years period from 2016 to 2021.

Sample size

All reported malaria cases at districts which had only primary metrological station in Amhara region from January 1, 2016 to December 30, 2021 were included in the analysis. A total of 2448 observation were included in the follow-up.

Data collection

Data was collected from Amhara regional state public health institute's public health emergency management directorate, Malaria elimination program team, and from Amhara Meteorological agency of 34 districts that have primary metrological stations.

Data analysis

Binary logistic regression analysis was applied to results from malaria outbreak based on Socio-demographic variables (age, location and migration), climatic factors (maximum and minimum temperature, relative humidity, rainfall, wind speed, sunshine), type of malaria species (*PF* and *PV)*, geographical factors (irrigation and altitude), seasonal adjusted factors and malaria outbreak were analyzed. We developed a mathematical model by binary logistic regression method. Based on our study malaria outbreak was defined as the incidence of malaria cases exceed the third quartile or second largest value of five years' previous data. Likewise when we had less than five years' data, we can say that any number of malaria cases more than double the number in the same month of last year's data was called an outbreak [10].

In the situation, when we have n predictor variable, the general regression model (Logit (p) = ln p/(1-p) = β0 + β1X1 + β2X2 +……+ βnXn Where: "p" is P (y = 1| X1, X2, X3…⋯Xn) and the probability that artifacts affected the predictor variable X1, X2, X3……Xn) are taken into consideration.

For this study, Missing data was assumed as missing at random [11] and was handled using by multiple imputation with chained equations (MICE) [12]. Sensitivity analysis was done to check whether the assumption of missing at random is valid or not. Variables with p-values of ≤ 0.25 were entered to the multivariable logistic regression model.

Backward stepwise multivariable logistic regression analysis was used to an elimination of not significant parameters and identifies predictors included into the final model. In this way, we could visualize the values of the coefficient β for all predictor variables, the elimination condition from the model for not statistically significant parameters, and the way in which the elimination of some parameters could affect the values of β coefficient and the mathematical model precision.

All models were developed at three time point with current month data, lag 1 month data (Model 1. lag 1) and lag 2-month data (Model 1. lag 2). Secondly, we included predictor variables and seasonal adjusted factors (SAF) into the model (Model 2). Seasonal decomposition was used to decompose the time series into a seasonal component, a combined trend and cycle component, as well as an error component [13].

To control for the impact of seasonality, we decomposed the malaria incidence into three series. That is, Yt = Tt + St + Et, where Yt denotes the malaria incidence, Tt denotes the trend component, St denotes the seasonal component, and Et denotes the residual component. To control the impact of seasonality in logistic regression model, we input the St into the logistic regression model as a seasonal factor, which indicate the effect of each period on the level of the series were used to determine the peak of seasonal variation [14]. Forecasting of the outbreak of monthly malaria was also done using the best fit model.

Predictors that had association with the malaria outbreak in the final model were reported using their beta coefficients with 95% CIs, and odds ratios (ORs) and risk scores. We tested the goodness-of-fit of model via maximum likelihood. The model accuracy was assessed by computing discrimination (area under the receiver operating characteristic curve (AUC), calibration (by calibration plot and Hosmer-Leme show model goodness of fit test) using "classifier plots" and "givitiR" packages of R respectively [5,15]. An AUC value ranges from 0.5 (no predictive ability) to 1 (perfect discrimination) [16,17]. The model was assumed to be well fitted when calibration test p-value is greater than 0.05 [18].

The regression beta coefficients, with its 95% confidence levels, and the AUC were adjusted for over fitting or optimism using bootstrapping technique. Internal validation for the model was performed by boot strapping [19] method which can be calculated by bootstrapping 1000 samples with replacement. The boot strapped regression coefficients and the AUC is considered as a predictive performance of the model that can be expected when the model is applied to similar populations in the future. Statistical analyses were performed using Statistical Package for the Social Science (SPSS) software version 25.0 and R statistical programming language version 4.0.4 with "Regression" and "Forecasting" procedures. Finally, result is presented by using frequencies, proportions, graph and tables.

Variables

Dependent variable: Malaria outbreak (Yes/No)

Independent variables: Location, Migration, Maximum and minimum temperature (°C), Relative humidity (%), Rainfall (mm), Wind speed (km/h), Sunshine, Altitude [11], Seasonality factors, Irrigation (m^3/s), *PF* by Blood film and RDT, *PV* by Blood film and RDT are the independent variables.

Operational definition

The forecasting model:

Malaria outbreak = $\alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$

Pr (Malaria outbreak) = exp/(1+exp(X)), where: X = $\beta_1 x_1 + \ldots + \beta_4 x_9$ [6]

Malaria outbreak: -when we had five years' previous data and malaria cases exceeded the third quartile number or second threshold value (or line on the chart) then there was an outbreak for that month likewise.

When we had less than five years of data, we can say that any number of malaria cases more than double the number in the same month of last year's data was called an outbreak.

Malaria case: Occurrence of malaria infection in a person in whom the presence of malaria parasites in the blood has been confirmed by a diagnostic test [20].

Climate zone

- **Kolla** (Tropical zone): Districts that had below 1830 meters in elevation
- **Woyna Dega** (Subtropical zone): Includes the highlands areas of 1830-2440 meters in elevation.
- **Dega** (Cool zone): Districts that had above 2440 meters in elevation.

Seasons

- **Bega (winter)** - June, July and August are the summer season. Heavy rainfalls in these three months.
- **Belg (autumn)** - September, October and November are autumn season sometime known as the harvest season.
- **Kiremt or Meher (summer)** - December, January and February are the dry season with frost in morning especially in January.
- **Tseday (spring)** - March, April and May are the autumn season with occasional showers. May is the hottest month in Ethiopia.

Results

A total of 34 (94%) primary metrological stations and nine districts with developmental corridors were included within 2448 observations. A total of 3,847,918 parasitological tests were performed by Microscopy and rapid diagnostic test (RDT) with the study period. Out of this total tested 919,586 malaria confirmed cases were reported with a test positivity rate of 23.9%. *Plasmodium falciparum* by Blood film and RDT accounted 532,298 (57.9%). Irrigation in the selected metrological study area accounted 368 (15%) (Table 1).

Development of empirical model and forecasting of Malaria outbreak

For the development of empirical model and forecasting of malaria outbreak, variable which have a p values of < 0.25 in bivariable analysis were nominated for multivariable logistic regression analysis. The multivariable logistic regression analysis result showed that the monthly outbreaks of malaria were significantly associated with rainfall, sunshine, irrigation and minimum temperature (Table 1).

The results of model 1 and model 2 showed that the monthly outbreaks of malaria were significantly

Table 1: Frequency and category of independent variables with Bivariable and Multivariable logistic regression analysis for the forecasting of malaria outbreak at Amhara region, Ethiopia, 2022.

Variables	Category	Number	(%)	Malaria outbreak		Bivariable analysis			Multivariable analysis			
				Yes	No	B	95 % CI	P value	Original β (95 % CI)	Bootstrap β	P-value	Simplified Risk score
Rainfall (mm)	RF < 98.18	1199	49	445	1990		**1**			**1**		
	RF > = 98.18	1249	51	11	2	3.203	(5.4-111.35)	< 0.001*	2.913 (4.189,100.55)	1.461	< 0.0001	63.5
Relative humidity (%)	RH < 52.67	2443	99.8	187	884		**1**					
	HR > = 52.67	5	0.2	269	1108	0.138	(0.122-9.796)	0.29				
Wind speed (m/s)	WS < 2.14	1224	50	456	1992		**1**					
	WS > = 2.14	1224	50	456	1992	- 0.022	(0.798-1.20)	0.836				
Sunshine	SH < 7.17	1525	62.3	445	1981		**1**			**1**		
	SH > = 7.17	923	37.7	11	11	1.493	(1.918-10.33)	< 0.001*	1.467 (1.621-9.146)	1.408	< 0.001	61
Irrigation	Absent	2080	85	409	1671		**1**		1			
	Present	368	15	47	321	-0.512	(0.433-0.829)	< 0.001*	-1.106 (1.161-2.420)	0.418	< 0.0001	18
Migration	Absent	2080	85	409	1671		**1**					
	Present	368	15	47	321	-0.512	(0.433-0.829)	< 0.001*				
Altitude	Dega	864	35.3	165	699		**1**					
	Woynadega	720	29.4	147	573	0.089	(0.853-1.402)	0.481				
	Kolla	864	35.3	144	720	-0.158	(0.67-1.093)	0.209				
Month	Summer	612	25.0	131	481		**1**					
	Spring	612	25.0	99	513	-0.2027	(0.595-1.049)	0.175				
	Winter	612	25.0	109	503	-0.094	(0.567-1.147)	0.542				
	Outman	612	25.0	117	495	-0.87	(0.687-1.224)	0.917				
Minimum Temperature (°C)				0	456		**1**					
				0	1992	-0.03	(0.951-0.991)	< 0.001				
Maximum Temperature (°C)				0	456		**1**					
				0	1992	-0.003	(0.982-1.023)	0.024	- 0.041(0.942, 0.990)	- 0.023	< 0.001	1

Bold with* = Nominated variable to multivariable logistic regression analysis at a p-value of < 0.25; Reference category is indicated by bold number 1; *Simplified risk score: Obtained by dividing the bootstrapped coefficients of predictors included in the reduced model by the smallest boot strapped coefficient (0.023); Minimum and maximum temperature = continues variable that was included in the study

Table 2: Malaria outbreak forecast models in, Amhara region, Ethiopia, 2022.

Models	Variable	B	95% CI	2 Log likelihood
Model 1	Irrigation area	-1.092	1.161-2.42	-2262.439
	Minimum temperature (°C)	-0.041	0.956-0.997	
	Rainfall (mm)	2.913	2.913-5.755	141.326
	Sunshine	1.467	1.711-9.791	
Model Lag 1	Rainfall (mm)	1.067	1.445-5.85	-2325.497
	Sunshine	-1.067	0.19-0.628	
Model Lag 2	Rainfall (mm)	1.058	1.423-5.837	-2342.813
Model 2	Minimum temperature (°C)	-0.041	0.934-0.983	
	Rainfall (mm)	2.913	4.189-141.326	
	Sunshine	1.467	1.776-10.582	

associated with rainfall, sunshine, irrigation and minimum temperature. The results of **model 1** showed that the monthly outbreaks of malaria were significantly associated with rainfall, sunshine, irrigation and minimum temperature. The predictors of minimum temperature increased by 1 unit (1 °C), the possibility of malaria outbreaks decreased by 2.3% (OR = 0.977, (0.956, 0.9979)) (Table 2).

The minimum temperature increased by one unit (1 °C); the possibility of malaria outbreaks decreased by 2.3 (OR = 0.977, 95% CI; 0.956-0.997). Areas that had sunshine exposer above the mean (≥ 7.167) had seven-fold higher odds of having monthly out breaks of malaria (OR = 7.10, (1.7106-9.791)) compared to areas with sunshine with above the mean (≥ 7.167).

The odds of having in areas with rainfall above the mean (≥ 98.178) were 21 times that of areas with rainfall below monthly rainfall (OR = 21.73, 95% CI; 5.755-141.32). Irrigation area had 1.52 times the odds of having monthly outbreaks of malaria compare to areas had not irrigation (OR = 1.52, 95% CI; 1.16-2.42) (Table 2).

The results of **model lag1** showed that the one month lagged effects on malaria outbreak has association with variables includes sunshine and rainfall. The odds of having areas with rain full above the mean (≥ 98.178) were one time that of areas with rainfall below monthly rainfall (OR = 1.067, (1.445-5.850)) (Table 2).

The results of **model lag 2** demonstrated that the lagged effects on malaria outbreak after two months later have association with rainfall. Areas that had sunshine exposer above the mean (≥ 7.167) had two-fold higher odds of having monthly out breaks of malaria (OR = 2.88, (0.190-0.628)) compared to areas with sunshine with above the mean (≥ 7.167) (Table 2).

Areas that had sunshine exposer above the mean (≥ 7.167) had seven-fold higher odds of having monthly out breaks of malaria (OR = 4.10, (1.7106, 9.791)) compared to areas with sunshine with above the mean (≥ 7.167). The odds of having in areas with rain full above the mean (≥ 98.178) were 21 time that of areas with rainfall below monthly rainfall (OR = 21.73, (5.755, 141.32)). Areas that had irrigation had 1.52 times the odds of having monthly outbreaks of malaria compare to areas had no irrigation (OR = 1.52, 1.16, 2.42) (Table 2).

After adjustment of seasonality as confounder, model 2 showed that there were significantly associations between the variables such irrigation (present), rainfall (≥ 98.178), sunshine (≥ 7.167), minimum temperature with monthly outbreaks of malaria (Table 2).

The AUC of the final original reduced model was 55.3% (95% confidence interval: 0.522-0.582). The calibration performance (fit of observed to expected risk a cross all individuals) assessed by the calibration belt using "givitiR Standardization Belt" in R programming had a p-value of 0.50 (Figure 1).

The original reduced model coefficients were internally validated by bootstrapping technique which gives an optimism coefficient of 1.618. This indicated that the original reduced model was over fitted (will overestimate the effect when applied to new external population in the future) and too much optimistic due to optimism coefficient is greater than 10%.

The forecasting model

Malaria outbreak = $\alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$

= -1.465 + 0.432* irrigation + 1.296* rainfall above monthly mean (≥ 98.174) + 1.464* sunshine above monthly mean (≥ 7.167) - 0.041 Tmin

Pr(Malaria outbreak) = exp [6]/(1+exp(X)), where: X = $\beta_1 x_1 + \ldots + \beta_4 x_9$

Calibration plot of the risk prediction model for malaria outbreak among malaria cases is presented in Figure 2.

Using *"Optimal Cutpoints" package* and *"kappa index"*, the model had sensitivity of 99.12%, specificity of 3.72%, and positive predictive value of 19.07% and, negative predictive value of 94.87% respectively. A decision curve analysis was performed using *"rmda*

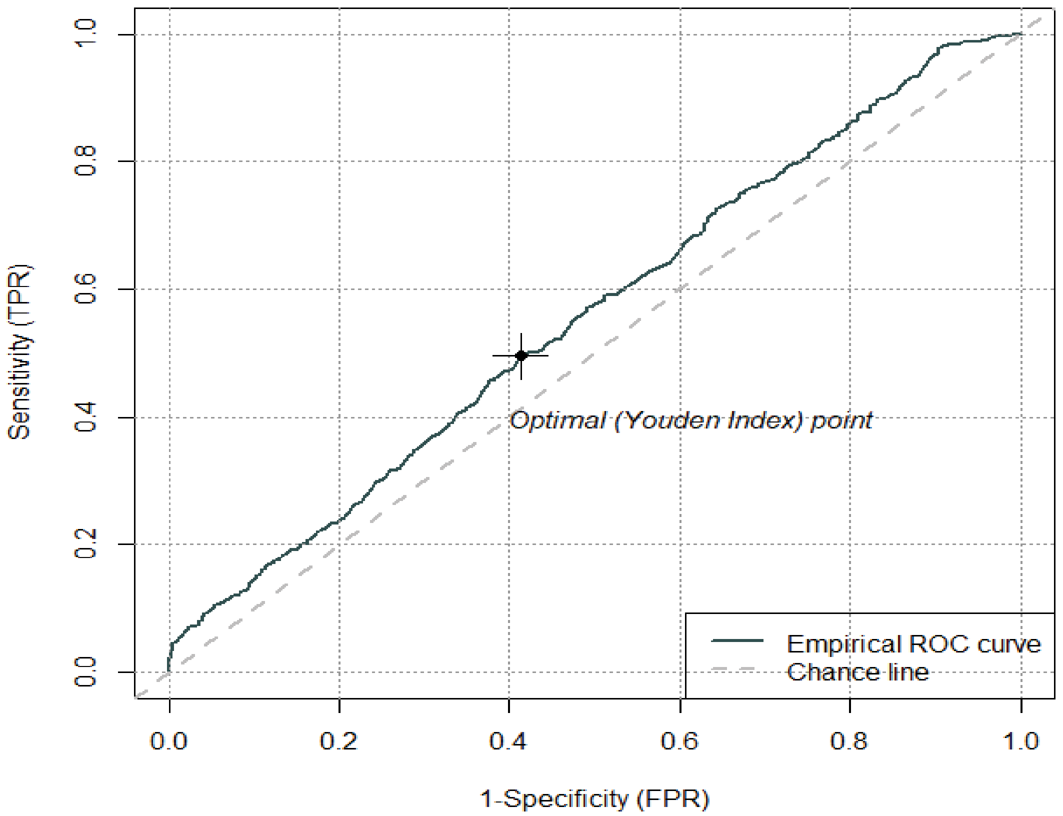

Figure 1: ROC_AUC of original risk prediction model for malaria outbreak among reported malaria cases at specified districts, Amhara region, Ethiopia, 2022.

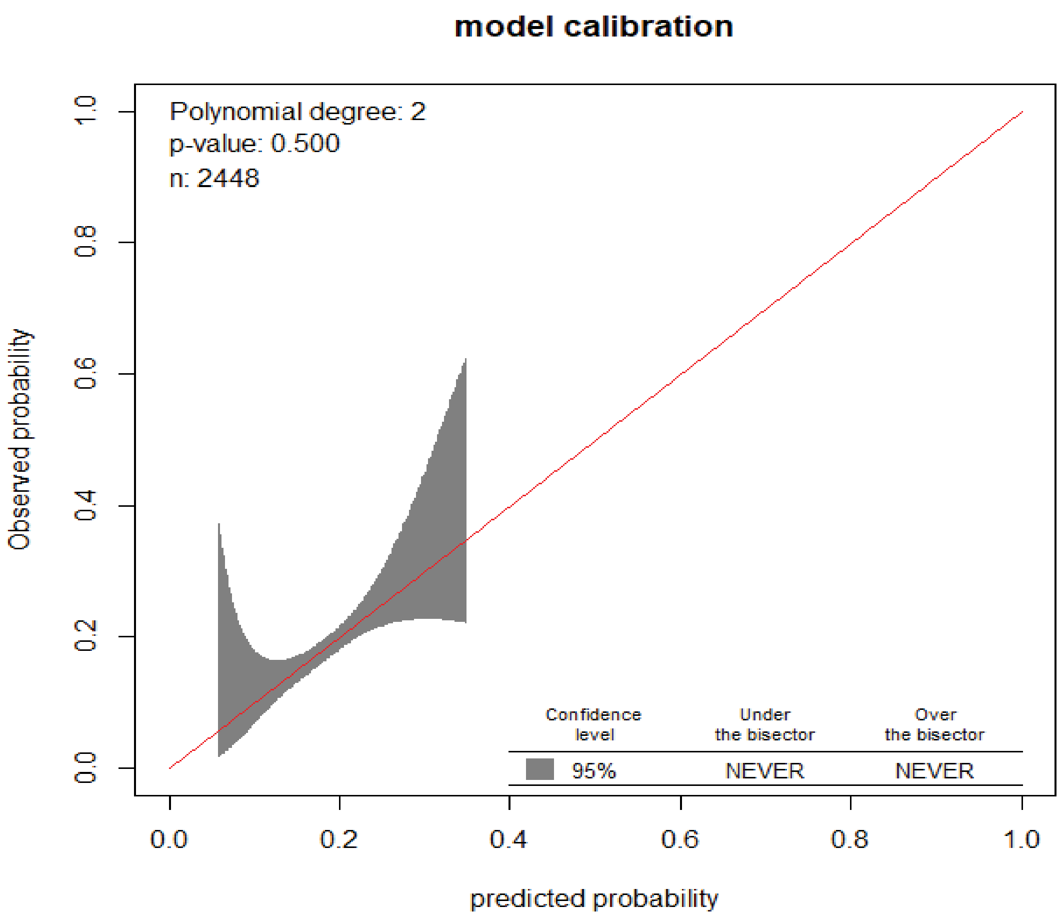

Figure 2: Calibration plot of the risk prediction model for malaria outbreak among malaria cases at specified districts, Amhara region, Ethiopia, 2022.

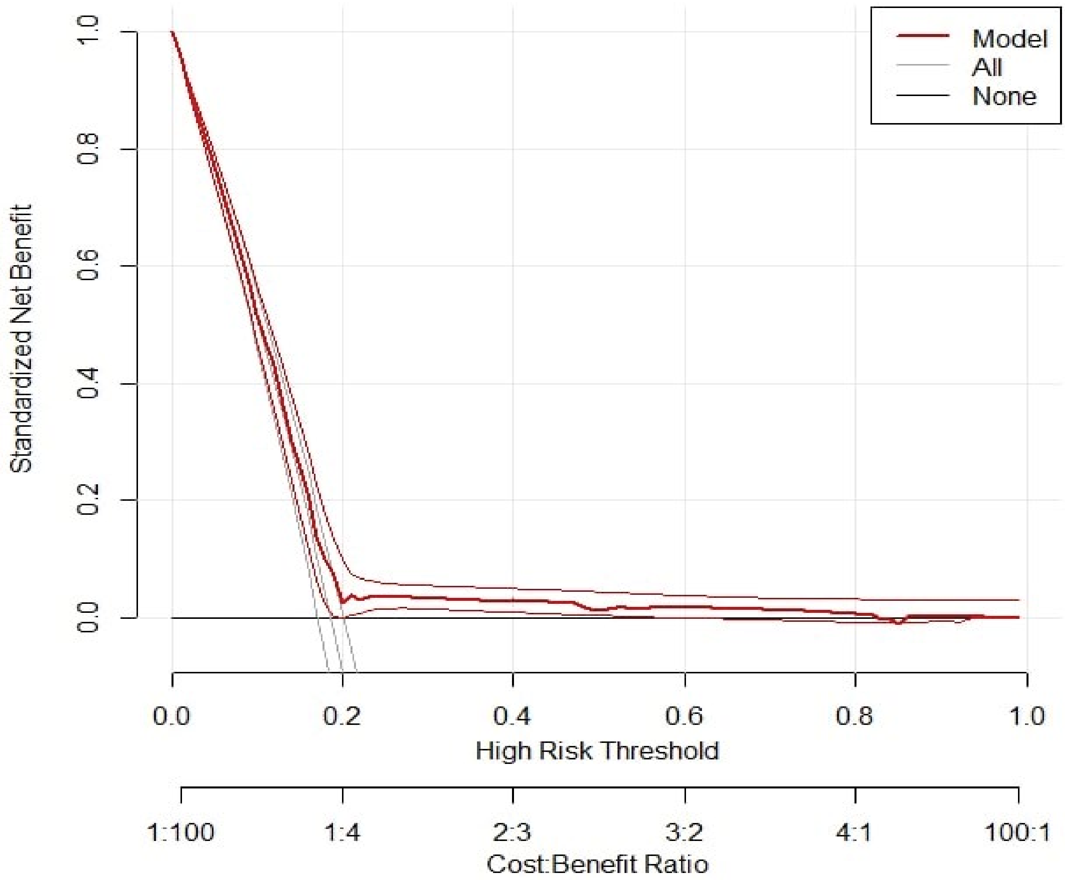

Figure 3: A decision curve plotting net benefit of the model against threshold probability of malaria outbreak among malaria cases at specified districts, Amhara region, Ethiopia, 2022.

package", and "*remotes package*" to evaluate the use of the prediction model across range of threshold probabilities, hence to assess clinical "net benefit" for the prediction model in comparison to default strategies. The model has the better net benefit across the entire range of threshold probabilities (Figure 3).

This model showed that predicting malaria outbreak at districts in the region for preparedness and provides intervention is better than giving intervention without using this model. Therefore, management decisions made using the model would provide a better net benefit to malaria outbreak. Generally, decisions made based on this model would have higher public health importance.

Discussion

In this study, current climate data was better to forecast monthly outbreaks of malaria. Sunshine, minimum temperature, rainfall and irrigation were optimal combinations to forecast malaria outbreak. The result of this finding is similar with a malaria forecasting model based on monthly case reports and climate variables conducted in Hefei, China [6]. This similarity may be due to the nature of data that we incorporate in the model.

This study simulated the effects of current lag 1, lag 2 and current data with seasonal decomposition on irrigation (present), rainfall above the mean(> = 98.178), sunshine above the mean (> = 7.167), minimum temperature on outbreaks of malaria within the study period at 34 districts/stations in Amhara region. **Model 1** with irrigation (present), rainfall (> = 98.178), sunshine (> = 7.167), and minimum temperature after adjustment of seasonal factors had best predictive effect. This study identified outbreaks of malaria associated with climatic factors, such as temperature, rainfall and sunshine and also geographical factors such as irrigation in Amhara region.

Malaria outbreaks appeared to be forecasted using the current climatic data, with specificity and sensitivity (SE: 99.12%, SP: 3.715%) even after adjustment for seasonality during January, 2016-December, 2021 in Amhara region. This is similar with a malaria forecasting model based on monthly case reports and climate variables in a result finding in Hefei, China [6]. This study provides evidence to support the meteorology and irrigation-malaria relationship and to determine the validity of meteorological and geographical conditions such as irrigation to create a general forecasting model of malaria outbreak.

Malaria outbreak was affected by sunshine, minimum temperature, rainfall and irrigation in this study. However, study done in China showed that, relative humidity, sunshine, and barometric pressure were significantly associated with malaria outbreaks after adjustment for seasonality [6]. This difference may be due to data that we incorporate in the model was collected from the station sites that was constructed across the road highway rather than malaria areas and also may be the difference of climatic condition between countries.

In this study, rainfall and temperature were significantly associated with malaria transmission. A study by Midekisa, et al. showed that rainfall rather than temperature is the most important predictor of malaria transmission in Africa [21,22]. However, a study conducted in Nigeria indicates that the association between malaria and temperature varied even in a country [23]. The discrepancy may be geographical difference and different drivers for malaria transmission in different ecological areas.

Sensitivity of the current model was in line with a study done on malaria early detection algorism in Amhara region by Nekorchuk, et al. (sensitivity = 80%-100%) [24]. The possible explanation for this similarity may be die to similarities in the study setting, climatic and geographic factors.

Sensitivity of the current model was higher than a study in East Africa [11] and in China [6]. Possible explanation may be due to differences in the prevalence of disease. It may also be from differences of predictors included in the study. Moreover, the model had high sensitivity (99.12%) it is less likely to predict false negative monthly outbreaks of malaria.

Specificity and positive predictive power of this model was lower than a malaria epidemic prediction models in East Africa by Githeko AK, Ogallo L, Lemnge M, Okia M, Ototo EN. using temperature and rainfall that had 75%, specificity of 99% and positive predictive power of 86% [25] and in Madagascar [26]**.** The reason may be from difference in data used for analysis in which a model Githeko AK, Ogallo L, Lemnge M, Okia M, Ototo EN., used malaria data only from hospitals (confirmed inpatient malaria data). Furthermore, the gap may be from differences in the quality of malaria and meteorological data (temperature and rainfall was collected from meteorological stations closest to the source of the malaria data) in the East African study. The other possible explanation may be the higher sensitivity of the current model that may decrease its specificity.

The specificity of the present model was lower than a furcating model done using Relative humidity sunshine and barometric pressure in an empirical model developed by Zhai JX, et al. in Hefei, China, 1990-2011 with a sensitivity of 70.52% and specificity of 70.30% [6]. Therefore, the present model will be more likely to forecast false negative monthly outbreaks of malaria. The gap in the specificity may be from differences in the predictors included in the model. The second reason maybe climate difference. Thirdly, the difference may be geographical difference between the countries.

In this study, minimum temperature, sunshine and rainfall above the mean and presence of irrigation were significantly associated with malarial outbreak in the study area. Whereas the finding was different from study done at Boricha District in the Simada regional state of Ethiopia [27]. The difference in study result might be due to meteorological station area coverage and recording difference in between the two areas and it needs further investigation.

The forecasting model developed in Ghana indicated that mean minimum and maximum monthly temperatures lagged at three months were significant interpreter of malaria incidence while rainfall was not [28]. This difference might be related to climatic and geographical difference of Ghana and Ethiopia.

Strength and Limitation of the Study

- Forecasting model was constructed using easily obtainable climatic variables such as temperature, rainfall, sunshine and geographical variables like irrigation.
- Metrological data is not restricted to selected districts rather it across roads and highways.

Migration and irrigation data were had information only in nine development corridors rather than specific registered number in the region

Conclusions

The incidence of monthly malaria outbreak was 18.62% in study period. A combination of predictors; sunshine, minimum temperature, rainfall and irrigation were important forecasting for malaria outbreak. Prediction model had poor discrimination power, not internally valid, but calibrated.

Recommendations

The recommendation would be the following.

For Researchers

- To conduct studies by using onsite recorded climatic data, irrigation distance related to malaria area, migrants related to its origin and other climatic related factors.

For Amhara Public health institute and other health agencies

- To apply the forecast model for forecasting monthly outbreaks of malaria
- To apply the forecast model in the malaria epidemic preparedness and response

For policy makers

- Consider this empirical forecasting model while developing malaria outbreak forecasting, warning and preparedness plans.

Decelerations

Ethics

Ethical clearance was obtained from the Ethical Review Board (IRB) of Bahir Dar University with an issue number 003 and date May 13, 2022. Besides this Permission letter was obtained from the Amhara public health institute, Research, and technology transform directorate with an issue number of APHI /03/1463 and date May 29, 2022 to the public health emergency management directorate which was located in the Amhara region. We have no direct contacts to participants during data collection period and do not recruit human participants for this study. Information that could identify individual participants during or after data collection had no access. Informed consent was waived by IRB of Bahir Dar University. The purposes and the importance of the study were stated to Amhara public health institute public health emergency management directorate malaria elimination program team.

Authors contribution

Conceptualization: Fetelework Workneh, Zelalem Mehari, Worku Awoke, Damtie Lankir; Data curation: Fetlework Workineh; Formal analysis: Fetlework Workineh, Zelalem Mehari, Worku Awoke, Damtie Lankir; Investigators: Fetlework Workineh, Zelalem Mehari, Worku Awoke; Methodology: Fetlework Workineh, Zelalem Mehari, Worku Awoke, Damtie Lankir; Supervision: Zelalem Mehari, Worku Awoke, Damtie Lankir; Visualization: Zelalem Mehari, Worku Awoke; Writing - original draft: Fetlework Workineh; Writing - review & editing: Damtie Lankir.

Data availability statement

All relevant data are within the paper.

Funding

This work was not supported by any funds.

Competing interests

No competing interest is present.

Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Patient consent

No direct contact from the patients.

Provenance and peer review

Not commissioned; externally peer reviewed.

Acknowledgements

The author would like to thanks Amhara Region Public Health Institute, Public Health Emergency Management Directorate and Amhara Region Metrological Agency.

References

1. Kiszewski A, Mellinger A, Spielman A, Malaney P, Sachs SE, et al. (2004) A global index representing the stability of malaria transmission. Am J Trop Med Hyg 70: 486-498.
2. Mace KE, Lucchi NW, Tan KR (2021) Malaria surveillance - United States, 2017. MMWR Surveill Summ 70: 1-35.
3. WHO (2021) WHO Guidelines for malaria.
4. Midekisa A, Senay GB, Wimberly MC (2014) Multisensor earth observations to characterize wetlands and malaria epidemiology in Ethiopia. Water Resource Research 50: 8791-8806.
5. Hassen HY, Gebreyesus SH, Endris BS, Roro MA, Van Geertruyden JP (2020) Development and validation of a risk score to predict low birth weight using characteristics of the mother: Analysis from BUNMAP cohort in Ethiopia. J Clin Med 9: 1587.
6. Zhai JX, Lu Q, Hu WB, Tong SL, Wang B, et al. (2018) Development of an empirical model to predict malaria outbreaks based on monthly case reports and climate variables in Hefei, China, 1990-2011. Acta Trop 178: 148-154.
7. Badr Bin Ashoor SWH (2019) Current Trends and Future Developments on (Bio-) Membranes.
8. Midekisa A, Beyene B, Mihretie A, Bayabil E, Wimberly MC (2015) Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. Parasit Vectors 8: 339.
9. Naish S, Mengersen K, Hu W, Tong S (2013) Forecasting the future risk of Barmah Forest virus disease under climate change scenarios in Queensland, Australia. PLoS One 8: e62843.
10. Midekisa A, Senay G, Henebry GM, Semuniguse P, Wimberly MC (2012) Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. Malar J 11: 165.
11. Githeko AK, Ogallo L, Lemnge M, Okia M, Ototo EN (2014) Development and validation of climate and ecosystem-based early malaria epidemic prediction models in East Africa. Malar J 13: 329.
12. Jakobsen JC, Gluud C, Wetterslev J, Winkel P (2017) When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 17: 162.
13. Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990) STL: A seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics 6: 3-73.
14. Kumar V, Singh A, Adhikary M, Daral S, Khokhar A, et al. (2014) Seasonality of tuberculosis in Delhi, India: A time series analysis. Tuberc Res Treat 2014: 514093.
15. Finazzi, S, Poole D, Luciani D, Cogo PE, Bertolini G (2011) Calibration belt for quality-of-care assessment based on dichotomous outcomes. PLoS One 6: e16110.
16. Grobbee DE, Hoes AW (2014) Clinical epidemiology:

Principles, methods, and applications for clinical research. (2nd edn), Jones & Bartlett Publishers.

17. Yang S, Berdine G (2017) The receiver operating characteristic (ROC) curve. The Southwest Respiratory and Critical Care Chronicles 5: 34-36.
18. Hosmer DW, Jovanovic B, Lemeshow S (1989) Best subsets logistic regression. Biometrics 45: 1265-1270.
19. Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, et al. (2012) Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio) marker. Heart 98: 683-690.
20. WHO (2021) WHO malaria terminology.
21. Nath DC, Mwchahary DD (2013) Association between climatic variables and malaria incidence: A study in Kokrajhar district of Assam, India. Global J Health Sci 5: 90-106.
22. Nigussie TZ, Zewotir TT, Muluneh EK (2022) Detection of temporal, spatial and spatiotemporal clustering of malaria in northwest Ethiopia, 2012-2020. Scientific Reports 12: 1-11.
23. Lee E, Burkhart J, Olson S, Billings AA, Patz JA, et al. (2016) Relationships of climate and irrigation factors with malaria parasite incidences in two climatically dissimilar regions in India. Journal of Arid Environments 124: 214-224.
24. Merkord CL, Liu Y, Mihretie A, Gebrehiwot T, Awoke W, et al. (2017) Integrating malaria surveillance with climate data for outbreak detection and forecasting: The EPIDEMIA system. Malar J 16: 89.
25. Sreekanth B, Shenoy S, Lella KS, Girish N, Reddy RS (2011) Evaluation of Blood Smears, Quantitative Buffy Coat and Rapid Diagnostic Tests in the Diagnosis of Malaria. J Bacteriol Parasitol 2: 8.
26. Mangani C, Frake AN, Chipula G, Mkwaila W, Kakota T, et al. (2022) Proximity of Residence to Irrigation Determines Malaria Risk and Anopheles Abundance at an Irrigated Agroecosystem in Malawi. Am J Trop Med Hyg 106: 283-292.
27. Srimath-Tirumula-Peddinti RCPK, Neelapu NRR, Sidagam N (2015) Association of climatic variability, vector population and malarial disease in district of Visakhapatnam, India: A modeling and prediction analysis. PLoS One 10: e0128377.
28. Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, et al. (2013) Born too soon: The global epidemiology of 15 million preterm births. Reprod Health 10: S2.