Analysis of Prevalence of Malaria and Anemia Using
Bivariate Probit Model

Senayit Seyoum

Abstract  Malaria and anemia are public health problems that have an impact on social and economic development. Malaria causes 70,000 deaths each year and accounts for 17% of outpatient visits to health institutions. It is one of the causes of anemia. Therefore, knowing the relation between malaria and anemia could have a great contribution to the development of prevention strategies. This study is intended to jointly model the prevalence of malaria and anemia by employing a bivariate probit model and show their relationship. The data was obtained from 384 patients visiting Alaba health center. The results of the bivariate probit model shows that sex, age, education level and marital status are significantly associated with malaria and sex and education level are significantly associated with anemia. The results of the seemingly unrelated bivariate probit model shows that sex, education level, age and marital status are significantly determining the prevalence of malaria, and malaria, sex and education level are significantly determining the prevalence of anemia.

Keywords  Malaria · Anemia · Bivariate probit · Seemingly unrelated bivariate probit

1 Introduction

Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected female Anopheles mosquitoes. About 3.2 billion people, almost half of the world’s population, are at risk of malaria [1]. According to 2015 WHO report, there were 214 million new cases of malaria worldwide in 2015 (range
The African Region accounted for most global cases of malaria (88%), followed by the South-East Asia Region (10%) and the Eastern Mediterranean Region (2%) [2].

It has been estimated that approximately 68% of the Ethiopian population live in areas below 2000 m of altitude and, thus, are considered to be at risk of malaria. This makes malaria the number one health problem in Ethiopia with an average of 5 million cases per year. Plasmodium falciparum and plasmodium vivax are the most dominant malaria parasites in Ethiopia [3,4]. The disease causes 70,000 deaths each year and accounts for 17% of outpatient visits to health institutions [5].

Anemia, defined as a low blood hemoglobin concentration, has been shown to be a public health problem that affects low, middle and high income countries and has significant adverse health consequences, as well as adverse impacts on social and economic development. Anemia may result from a number of causes, with the most significant contributor being iron deficiency. Approximately 50% of cases of anemia are considered to be due to iron deficiency, but the proportion probably varies among population groups and in different areas, according to the local conditions. Other causes of anemia include acute and chronic infections (e.g. malaria, cancer, tuberculosis and HIV) [6].

The malaria parasites, entering the blood after an infective mosquito bite, infect red blood cells. At the end of that infection cycle, red blood cell ruptures. This process lowers the amount of red blood cells and can in a severe stage cause severe anemia [1]. On the African continent, where Plasmodium falciparum is the most prevalent human malaria parasite, anemia is responsible for about half of the malaria-related deaths [7]. Therefore, knowing the relation between malaria and anemia could have a great contribution to the development of prevention strategies. Although different studies have been conducted on the prevalence of anemia and malaria and to identify the association between them, still not much have been done to jointly model the prevalence of malaria and anemia and show the relationship between malaria and anemia. So this study is intended to jointly model the prevalence of malaria and anemia by employing a bivariate probit model.

2 Methodology

2.1 Data

The study was conducted on patients visiting Alaba Health Center, Alaba District, Southern Ethiopia, for medical examination from November to December 2016. The data was obtained from 384 study participants.

2.2 Bivariate Probit Model (BPM)

The bivariate probit model is a joint model for two binary dependent variables whose disturbances are assumed to be correlated. It generalizes the index function model from one latent variable to two latent variables that may be correlated.
Let $Y_1^*$ and $Y_2^*$ be two latent variables. A latent variable is a variable that is incompletely observed. Latent variables can be introduced into binary outcome models as an index of an unobserved propensity for the event of interest to occur [8].

Define the unobserved latent variables

\begin{align}
Y_1^* &= X_1 \beta_1 + \varepsilon_1 \\
Y_2^* &= X_2 \beta_2 + \varepsilon_2
\end{align}

where $\varepsilon_1$ and $\varepsilon_2$ are joint normal with means zero, variances one, and correlation $\rho$. 

\[ \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} | X \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \]

Then the bivariate probit model specifies the observed outcomes to be

\begin{align}
Y_1 &= \begin{cases} 1 & \text{if } Y_1^* > 0 \\ 0, & \text{otherwise} \end{cases} \\
Y_2 &= \begin{cases} 1 & \text{if } Y_2^* > 0 \\ 0, & \text{otherwise} \end{cases}
\end{align}

The bivariate probit model can be written as

\[ P (y_1 = i, y_2 = j) = \Phi_2 \left( X_1' \beta_1, X_2' \beta_2, \rho \right) \]

2.2.1 Estimation of Parameter Coefficients

The coefficients ($\beta_1$, $\beta_2$ and $\rho$) can be estimated using the maximum likelihood estimation.

First let’s define some function: let $q_1 = 2y_1 - 1$ and $q_2 = 2y_2 - 1$. Thus, $q_i = 1$ if $y_i = 1$ and $-1$ if $y_i = 0$ for $i = 1 and 2$.

The probabilities that enter the likelihood function are

\[ P (Y_1 = y_i, Y_2 = y_j | X_1 X_2) = \Phi_2 \left( q_1 X_1' \beta_1, q_2 X_2' \beta_2, q_1 q_2 \rho \right) \]

The likelihood function is given as

\[ L = \prod P (Y_1 = y_i, Y_2 = y_j) = \prod \Phi_2 \left( q_1 X_1' \beta_1, q_2 X_2' \beta_2, q_1 q_2 \rho \right) \]

Consider the latent variables and the observed outcomes on Eqs. (1), (2), (4) and (5). From these equations we can get the following relations.

\begin{align}
Y_1 &= 1 \text{ If } \varepsilon_1 > -X_1 \beta_1 \text{ and } Y_1 = 0 \text{ if } \varepsilon_1 < -X_1 \beta_1 \\
\text{Also } Y_2 &= 1 \text{ if } \varepsilon_2 > -X_2 \beta_2 \text{ and } Y_2 = 0 \text{ if } \varepsilon_2 < -X_2 \beta_2
\end{align}
We can write the likelihood as

\[ L = \prod \left\{ P \left( \varepsilon_1 > -X_1\beta_1, \varepsilon_2 > -X_2\beta_2 \right) + P \left( \varepsilon_1 < -X_1\beta_1, \varepsilon_2 > -X_2\beta_2 \right) \\
+ P \left( \varepsilon_1 > -X_1\beta_1, \varepsilon_2 < -X_2\beta_2 \right) + P \left( \varepsilon_1 < -X_1\beta_1, \varepsilon_2 < -X_2\beta_2 \right) \right\} \]

(12)

Then the log likelihood is

\[ \ln L = \sum \ln P \left( Y_1 = y_i, Y_2 = y_j \right) = \sum \ln \Phi_2 \left( q_1 X_1' \beta_1, q_2 X_2' \beta_2, \rho \right) \]

(13)

\[ \ln L = \sum \left\{ \ln P \left( \varepsilon_1 > -X_1\beta_1, \varepsilon_2 > -X_2\beta_2 \right) + \ln P \left( \varepsilon_1 < -X_1\beta_1, \varepsilon_2 > -X_2\beta_2 \right) \\
+ \ln P \left( \varepsilon_1 > -X_1\beta_1, \varepsilon_2 < -X_2\beta_2 \right) + \ln P \left( \varepsilon_1 < -X_1\beta_1, \varepsilon_2 < -X_2\beta_2 \right) \right\} \]

(14)

\[ \ln L = \sum \left\{ \ln \Phi_2 \left( X_1' \beta_1, X_2' \beta_2, \rho \right) + \ln \Phi_2 \left( -X_1' \beta_1, X_2' \beta_2, -\rho \right) \\
+ \Phi_2 \left( X_1' \beta_1, -X_2' \beta_2, -\rho \right) + \ln \Phi_2 \left( -X_1' \beta_1, -X_2' \beta_2, \rho \right) \right\} \]

(15)

So we can get the estimator for the coefficients by maximizing the log likelihood function.

### 2.2.2 Seemingly Unrelated Bivariate Probit Model

In this study it is also considered the application of the seemingly unrelated bivariate probit model. Seemingly unrelated bivariate probit model is used when two equations are to be estimated and the dependent variable of one of them is an explanatory variable in the other.

The latent variable for the seemingly unrelated bivariate probit model is given as:

\[ Y_1^* = X_1\beta_1 + \varepsilon_1 \]

(16)

\[ Y_2^* = \gamma Y_1 + X_2\beta_2 + \varepsilon_2 \]

(17)

The estimation procedures are the same except including \( \gamma Y_1 \) in this model.

### 2.2.3 Marginal Effects

After we estimate the parameters we have to consider the marginal effects of the covariates in the conditional distribution [10, 11]. Marginal effects determine the magnitude of change of the conditional probability of the outcome variable when you change the value of a regressor, holding all the regressors constant at some value [9]. The marginal effect for the bivariate probit model is then given by

\[ \frac{\partial \Phi_2 \left( X_1' \beta_1, X_2' \beta_2, \rho \right)}{\partial X_i} = \varphi \left( X_i' \beta_i \right) \Phi_2 \left( \frac{X_i' \beta_2 - \rho X_i' \beta_1}{\sqrt{1 - \rho^2}} \right) \beta_i, i = 1, 2 \]

(18)

The marginal effects for categorical variables shows how conditional probability changes as the categorical variable changes from 0 to 1, after controlling in some way for the other variables in the model.
Table 1  The frequency (percentage) of each variable included in the study

| Variable   | Codes/codes          | Frequency (percentage) |
|------------|----------------------|------------------------|
| Malaria    | 0 = Absent 330 (85.94%) 1 = Present 54 (14.06%) |
| Anemia     | 0 = Absent 290 (75.52%) 1 = Present 94 (24.48%) |
| Sex        | 0 = Male 153 (39.84%) 1 = Female 231 (60.16%) |
| Age        | 0 = 15–30 years 68 (17.71%) 1 = 31–45 years 182 (47.4%) 2 = 46–60 years 108 (28.13%) 3 ≥ 60 years 26 (6.77%) |
| Education level | 0 = Illiterate 178 (46.35%) 1 = Read and write 64 (16.67%) 2 = Primary education 92 (23.96%) 3 = Secondary education 33 (8.59%) 4 = Higher education 17 (4.43%) |
| Marital status | 0 = Single 127 (33.07%) 1 = Married 232 (60.42%) 2 = Widowed 10 (2.60%) 3 = Divorced 15 (3.91%) |

3 Result and Discussion

3.1 Descriptive Statistics

At first a summary statistic was made for each of the variables in the study. The summary measure is important to get a picture about the variables included in this study. The frequency and the percentage of each variable are calculated. The result is shown in the Table 1.

3.2 Chi Square Test of Association

Before we fit the probit model the relationship between the response and the explanatory variables is checked using a chi-square test. Chi-square test is used to see the association between each response variable, i.e. malaria and anemia, and the explanatory variables so that we can see whether the selected variables have a relation with the response variables. The result of the chi-square test is then shown in the Table 2.

As can be seen on Table 2 the variables sex, age and education level are associated with malaria and the variables sex and education level are associated with anemia. Marital status is found to be not associated with both malaria and anemia and age is also not associated with anemia.
Table 2  Chi-square test of association between the responses (malaria and anemia) and each explanatory variable

| Explanatory variable | Malaria | Anemia |
|----------------------|---------|--------|
|                      | Chi-square | df | p value | Chi-square | df | p value |
| Sex                  | 6.4718  | 1  | 0.011** | 3.2647  | 1  | 0.071*  |
| Age                  | 12.3083 | 3  | 0.006***| 2.0588  | 3  | 0.560   |
| Education level      | 9.8137  | 4  | 0.044** | 14.1583 | 4  | 0.007***|
| Marital status       | 0.4773  | 3  | 0.924   | 0.3057  | 3  | 0.959   |

***Statistically significant at 1%, **statistically significant at 5% and *statistically significant at 10%

3.3 Univariate Analysis

A univariate probit regression is fitted for malaria and anemia to see the effect of each explanatory variable when fitted with the other explanatory variables. The results obtained from the univariate models are discussed in the next part. The first category of each variable is considered as a reference category.

3.3.1 Univariate Analysis for Malaria

The prevalence of malaria is fitted with the other explanatory variables to identify the factors that affect the prevalence of malaria. From the chi square test of association it was found that sex, age and education level are significant. Then all the variables are fitted using a probit model and the result obtained is shown in Table 3.

Table 3  The parameter estimates and standard errors of the univariate probit model for malaria

| Explanatory variable | Parameter estimate | Standard error | Z statistic | p value |
|----------------------|--------------------|----------------|-------------|---------|
| Sex                  |                    |                |             |         |
| Female               | −0.4299            | 0.1837         | −2.34       | 0.019***|
| Age                  |                    |                |             |         |
| 31–45                | −1.1395            | 0.3415         | −3.34       | 0.001***|
| 46–60                | −2.0005            | 0.4136         | −4.84       | 0.000***|
| > 60                 | −2.8569            | 0.7722         | −3.70       | 0.000***|
| Education            |                    |                |             |         |
| Read and write       | −0.3787            | 0.2436         | −1.55       | 0.120   |
| Primary education    | −1.2753            | 0.2984         | −4.27       | 0.000***|
| Secondary education  | −0.7889            | 0.3578         | −2.21       | 0.027** |
| Higher education     | −0.7218            | 0.5230         | −1.38       | 0.168   |
| Marital status       |                    |                |             |         |
| Married              | 0.6372             | 0.2910         | 2.19        | 0.029** |
| Widowed              | 0.8584             | 0.5567         | 1.54        | 0.123   |
| Divorced             | 1.4014             | 0.7111         | 1.97        | 0.049** |

***Statistically significant at 1%, **statistically significant at 5% and *statistically significant at 10%
Table 4  The parameter estimates and standard errors of the univariate probit model for anemia

| Explanatory variable | Parameter estimate | Standard error | Z statistic | p value  |
|----------------------|--------------------|----------------|-------------|----------|
| **Malaria**          |                    |                |             |          |
| Present              | 0.7798             | 0.2085         | 3.74        | 0.000*** |
| Sex                  |                    |                |             |          |
| Female               | 0.3899             | 0.1646         | 2.37        | 0.018**  |
| Age                  |                    |                |             |          |
| 31-45                | 0.3056             | 0.2898         | 1.05        | 0.292    |
| 46–60                | 0.1512             | 0.3437         | 0.44        | 0.660    |
| > 60                 | −0.1394            | 0.4925         | −0.28       | 0.777    |
| **Education**        |                    |                |             |          |
| Read and write       | −0.4731            | 0.2205         | −2.15       | 0.032**  |
| Primary education    | −0.2695            | 0.2144         | −1.26       | 0.209    |
| Secondary education  | −0.9202            | 0.3534         | −2.60       | 0.009*** |
| Higher education     | −1.1404            | 0.5133         | −2.22       | 0.026**  |
| **Marital status**   |                    |                |             |          |
| Married              | −0.2658            | 0.2349         | −1.13       | 0.258    |
| Widowed              | −0.3339            | 0.5149         | −0.65       | 0.517    |
| Divorced             | −0.3582            | 0.5112         | −0.70       | 0.483    |

***Statistically significant at 1%, **statistically significant at 5% and *statistically significant at 10%

From Table 3, the result shows that sex, age, education level and marital status are significantly determining the prevalence of malaria. From the test of association marital status was found to be insignificant but when fitted with other variables it becomes significant.

3.3.2 Univariate Analysis for Anemia

The prevalence of anemia is fitted with the other explanatory variables to identify the risk factors. From the chi square test of association it was found that sex and education level are significant. Then all the variables are fitted using a probit model and the result obtained is shown in Table 4.

From the above table showing the univariate analysis of anemia, the result shows that malaria, sex and education level are significantly determining anemia. The variables age and marital status are not significant. This result shows that sex and education are the variables that are significantly affecting the prevalence of anemia whether fitted alone or together with other variables.

3.4 Bivariate Analysis for Malaria and Anemia

A bivariate probit model is used to jointly model the prevalence of malaria and anemia and the result is shown in the table below. The data was modeled first using a bivariate
### Table 5  The parameter estimates and standard errors of the bivariate probit model for malaria and anemia

| Response variable | Explanatory variable | Parameter estimate | Standard error | Z statistic | p value | Marginal effects |
|-------------------|----------------------|--------------------|----------------|------------|---------|-----------------|
| Malaria           |                      |                    |                |            |         |                 |
|                   | Sex                  |                    |                |            |         |                 |
|                   | Female               | −0.4058            | 0.1821         | −2.23      | 0.026** | −0.0795         |
|                   | Age                  |                    |                |            |         |                 |
|                   | 31–45                | −1.1160            | 0.3361         | −3.32      | 0.001***| −0.3226         |
|                   | 46–60                | −1.9719            | 0.4103         | −4.81      | 0.000***| −0.4544         |
|                   | > 60                 | −2.6089            | 0.6835         | −3.82      | 0.000***| −0.4906         |
|                   | Education level      |                    |                |            |         |                 |
|                   | Read and Write       | −0.3830            | 0.2424         | −1.58      | 0.114   | −0.0912         |
|                   | Primary education    | −1.2881            | 0.2977         | −4.33      | 0.000***| −0.2116         |
|                   | Secondary education  | −0.7820            | 0.3563         | −2.2       | 0.028** | −0.1597         |
|                   | Higher education     | −0.7246            | 0.5144         | −1.41      | 0.159   | −0.1514         |
|                   | Marital status       |                    |                |            |         |                 |
|                   | Married              | 0.5841             | 0.2844         | 2.05       | 0.040** | 0.0980          |
|                   | Widowed              | 0.8772             | 0.5357         | 1.64       | 0.102   | 0.1671          |
|                   | Divorced             | 1.1611             | 0.6543         | 1.77       | 0.076*  | 0.2455          |
|                   |                      |                    |                |            |         |                 |
| Anemia            |                      |                    |                |            |         |                 |
|                   | Sex                  |                    |                |            |         |                 |
|                   | Female               | 0.2828             | 0.1581         | 1.79       | 0.074*  | 0.0815          |
|                   | Age                  |                    |                |            |         |                 |
|                   | 31–45                | 0.0928             | 0.2787         | 0.33       | 0.739   | 0.0290          |
|                   | 46–60                | −0.2068            | 0.3243         | −0.64      | 0.524   | −0.0590         |
|                   | > 60                 | −0.5939            | 0.4713         | −1.26      | 0.208   | −0.1458         |
|                   | Education level      |                    |                |            |         |                 |
|                   | Read and write       | −0.5238            | 0.2164         | −2.42      | 0.015** | −0.1637         |
|                   | Primary education    | −0.4683            | 0.2065         | −2.27      | 0.023** | −0.1491         |
|                   | Secondary education  | −1.0243            | 0.3448         | −2.97      | 0.003***| −0.2648         |
|                   | Higher education     | −1.2296            | 0.5038         | −2.44      | 0.015** | −0.2912         |
|                   | Marital status       |                    |                |            |         |                 |
|                   | Married              | −0.1385            | 0.2280         | −0.61      | 0.544   | −0.0413         |
|                   | Widowed              | −0.1448            | 0.4990         | −0.29      | 0.772   | −0.0431         |
|                   | Divorced             | −0.1295            | 0.4990         | −0.26      | 0.795   | −0.0387         |

\[ \hat{\rho} = 0.4014 \pm 0.1026 \]

***Statistically significant at 1%, **statistically significant at 5% and *statistically significant at 10%

...probit model and it was then fitted using the seemingly unrelated bivariate probit model. After that we can choose the model that best fits the data. Table 5 shows the bivariate analysis of malaria and anemia using the bivariate probit model.

As can be seen on Table 5 the result shows that sex, age, education level and marital status are significant in explaining the prevalence of malaria. This means these variables: sex, age, education level and marital status have an impact on the distribu-
tion of malaria disease. And sex and education level are significant in explaining the prevalence of anemia. This means having a different sex and education level have a difference on being affected by anemia. The average marginal effect was also calculated and the result shows that females have a 7.95% lower chance of being infected with malaria than males. People in age group of 31–45, 46–60 and above 60 have respectively a 32.26, 45.44 and 49.06% lower chance of being infected with malaria than that in age group of 15–30. People with primary education and secondary education have respectively 21.16 and 15.97% lower chance of being infected with malaria than that of illiterate people. Married and divorced people have respectively a 9.8 and 24.55% higher chance of being infected with malaria than that of single individuals. The average marginal effects for anemia also shows that females have 8.15% higher chance of being anemic than males. People who can read and write, having primary education, secondary education and higher education have respectively 16.37, 14.9, 26.5 and 29.12% lower chance of being infected with malaria than that of illiterate people.

The overall fit of the model is checked by the Wald test \([W^2 (22) = 63.72, p \text{ value} = 0.0000]\) and shows that the model is significant. It is also checked for the presence of correlation. The likelihood ratio test rejects the null hypothesis of zero correlation \([X^2 (1) = 12.8119, p \text{ value} = 0.0003]\). This means the correlation between the prevalence of malaria and anemia is significant.

The data was also modeled using a seemingly unrelated bivariate probit model. The result is shown in the Table 6. In this model we included only the significant variables for both the response variables.

As can be seen from Table 6, all the variables are significant for malaria. That mean all this variables; age, sex, education level and marital status, have an impact on the distribution of malaria disease. For anemia, malaria, sex and education are significant. That is; malaria, sex and education have an impact on the spread of anemia disease. The average marginal effect was also calculated and the result shows that females have 8.6% lower chance of being infected with malaria than males. People in age group of 31–45, 46–60 and above 60 have respectively a 30.44, 43.84 and 48.82% lower chance of being infected with malaria than that in age group of 15–30. People with primary education and secondary education have respectively 20.48 and 15.89% lower chance of being infected with malaria than that of illiterate people. Married and divorced people have respectively a 10.27 and 36.47% higher chance of being infected with malaria than that of single individuals. The average marginal effects of anemia shows that people with malaria have 46.62% higher chance of becoming anemic as compared to that of malaria free individuals. Females have 10.44% higher chance of being anemic than that of males. People who can read and write, having secondary education and higher education have respectively 11.17, 16.15 and 20.28% lower chance of being anemic than that of illiterate people.

The overall fit of the model is significant with Wald \([W^2 (23) = 75.28 \text{ and } p \text{ value} = 0.0000]\). The likelihood ratio test fails to reject the null hypothesis of zero correlation \([X^2 (1) = 0.8791, p \text{ value} = 0.3484]\). This means the correlation between the prevalence of malaria and anemia is not significant.
Table 6  The parameter estimates and standard errors of the seemingly unrelated bivariate probit model for malaria and anemia

| Response variable | Explanatory variable | Parameter estimate | Standard error | Z statistic | p value | Marginal effects |
|-------------------|----------------------|--------------------|----------------|------------|---------|-----------------|
| Malaria           | Sex                  |                    |                |            |         |                 |
|                   | Female               | -0.4390            | 0.1835         | -2.39      | 0.017**| -0.0857         |
|                   | Age                  |                    |                |            |         |                 |
|                   | 31–45                | -1.0631            | 0.3551         | -2.99      | 0.003***| -0.3044         |
|                   | 46–60                | -1.9198            | 0.4296         | -4.47      | 0.000***| -0.4384         |
|                   | > 60                 | -3.0576            | 0.8737         | -3.50      | 0.000***| -0.4882         |
|                   | Education level      |                    |                |            |         |                 |
|                   | Read and write       | -0.3728            | 0.2432         | -1.53      | 0.125   | -0.0882         |
|                   | Primary education    | -1.2326            | 0.3022         | -4.08      | 0.000***| -0.2048         |
|                   | Secondary education  | -0.7896            | 0.3540         | -2.23      | 0.026** | -0.1589         |
|                   | Higher education     | -0.7422            | 0.5281         | -1.41      | 0.160   | -0.1523         |
|                   | Marital status       |                    |                |            |         |                 |
|                   | Married              | 0.6152             | 0.2933         | 2.10       | 0.036** | 0.1130          |
|                   | Widowed              | 0.7542             | 0.5794         | 1.30       | 0.193   | 0.1404          |
|                   | Divorced             | 1.5697             | 0.7875         | 1.99       | 0.046** | 0.4266          |
| Anemia            | Present              | 1.3498             | 0.5914         | 2.28       | 0.022** | 0.4662          |
|                   | Sex                  |                    |                |            |         |                 |
|                   | Female               | 0.3986             | 0.1639         | 2.43       | 0.015** | 0.1044          |
|                   | Education level      |                    |                |            |         |                 |
|                   | Read and write       | -0.4089            | 0.2122         | -1.93      | 0.054*  | -0.1117         |
|                   | Primary education    | -0.1421            | 0.1921         | -0.74      | 0.459   | -0.0422         |
|                   | Secondary education  | -0.6423            | 0.3478         | -1.85      | 0.065*  | -0.1615         |
|                   | Higher education     | -0.8868            | 0.5150         | -1.72      | 0.085*  | -0.2028         |

\( \hat{\rho} \)

\( \hat{\rho} = -0.3492 \)

***Statistically significant at 1%, **statistically significant at 5% and *statistically significant at 10%

4 Discussion

The finding of this study shows that sex, age, education level and marital status are significant for malaria in both the univariate and the bivariate case. The study showed that females have the highest risk of malaria than males. Other study in Ethiopia also found the same result [12].

The result of this study also showed that there is a high risk of malaria among younger people than the older ones. This result is consistent with other studies. A study done in Ethiopia found that there is a high prevalence of malaria on younger age groups [12,13].

Based on the result found from our study, education level is found to be significantly associated with malaria. A study done in Kenya also found that greater malaria risk is
associated with lower level of education of individuals [14]. This study also showed that marital status is significantly associated with malaria. The result found that divorced people have the highest chance of being infected with malaria. This result is not consistent with other studies [15].

Malaria, sex and education level are found to be significantly associated with anemia in both the univariate and the bivariate models. According to this study people with malaria have the highest risk of being anemic than that of malaria free peoples. Another study in Ethiopia also found that malaria infected individuals have a high risk of being anemic than malaria free individuals [16].

Based on the result of this study females have the highest risk of anemia than males. A study done in south India also found that females have highest risk of anemia [17]. Another study in Ethiopia also found a consistent result [18]. Our finding also found that there is an association between anemia and education level. The result shows that illiterate people have the highest chance of anemia. Two studies done in India also found that people at lower education level have the highest risk of anemia [19,20].

5 Conclusion

First a chi-square test was conducted to see the association between the prevalence of malaria and anemia and the explanatory variables and the result shows that the variables sex, age and education level are associated with malaria and the variables sex and education level are associated with anemia. Then a univariate probit model was fitted for both malaria and anemia to see the effect of each explanatory variable when fitted with the other explanatory variables. The result shows that sex, age, education level and marital status are significant for malaria. Anemia is significantly associated with malaria, sex and education level.

A bivariate probit model was also fitted using both bivariate probit model and seemingly unrelated bivariate probit model. The results of the bivariate probit model shows that sex, age, education level and marital status are significantly associated with malaria and sex and education level are significantly associated with anemia. The average marginal effect also shows that males, people at the age of 15–30, illiterate people and divorced people have the highest chance of being infected with malaria and females and illiterate people have the highest chance of being anemic. The likelihood ratio test shows that there is a positive correlation between malaria and anemia. The result of the seemingly unrelated bivariate probit model shows that the variables significant for malaria are the same as that of the bivariate probit results but for anemia, malaria, sex and education level are significant. Based on the result of the average marginal effects it was found that males, people at the age of 15–30, illiterate people and divorced people have the highest chance of being infected with malaria. Females and illiterate people have the highest chance of being anemic and people with malaria have the highest chance of becoming anemic as compared with people who are not infected with malaria. The likelihood ratio test shows that there is no correlation between malaria and anemia.
References

1. WHO (2016) Fact sheet: Malaria. http://www.who.int/mediacentre/factsheets/fs094/en/. Accessed on 10 Mar 2017
2. WHO. World Malaria Report 2015. http://www.who.int/malaria/media/world-malaria-report-2015/en/. Accessed on 16 Mar 2017
3. National Malaria Control Team, Ethiopian Public Health Institute, World Health Organization, Addis Ababa University and the INFORM Project (2014) An epidemiological profile of malaria in Ethiopia. A report prepared for the Federal Ministry of Health, Ethiopia, the Roll Back Malaria Partnership and the Department for International Development, UK. March, 2014
4. Regasa B (2014) Magnitude of malaria infection in Ethiopia. Glob J Med Res 14(7–C):18–22
5. President’s Malaria Initiative (2008) Malaria operational plan (MOP) Ethiopia. FY 2008
6. WHO (2015) The global prevalence of anemia in 2011. World Health Organization, Geneva. http://www.who.int/nutrition/publications/micronutrients/global_prevalence_anemia_2011/en/. Accessed on 16 Mar 2017
7. Quintero JP, Siqueira AM, Tobón A, Blair S, Moreno A, Arévalo-Herrera M, Vinícius Guimarães Lacerda M, Valencia SH (2011) Malaria-related anemia: a Latin American perspective. Mem Inst Oswaldo Cruz, Rio de Janeiro 106(Suppl. 1):91–104
8. Colin Cameron A, Trivedi PK (2005) Micro econometrics: methods and applications. Cambridge University Press, Cambridge
9. Wambui Njoroge T, Musili Mwalili S, Kibira Wanjoya A (2015) A bivariate probit model for correlated binary data with application to HIV and male circumcision. Am J Theor Appl Stat 4(6):555–561. https://doi.org/10.11648/j.ajtas.20150406.27
10. Greene WH (1996) Marginal effects in the bivariate probit model. NYU Working Paper No. EC-96-11
11. Greene WH (2012) Econometric analysis, 7th edn. Prentice Hall Publishing, Upper Saddle River
12. Tefera G (2014) Prevalence of Malaria and associated factors among patients attending at Hallaba Health Center, Southern Ethiopia. Immunol Infect Dis 2(3):25–29
13. Zewotir TT, Mwambi HG (2013) The risk factor indicators of malaria in Ethiopia. Int J Med Med Sci 5(7):335–347
14. Ernst KC, Lindblade KA, Koech D, Sumba PO, Kuwuor DO, John CC, Wilson ML (2009) Environmental, socio-demographic and behavioral determinants of malaria risk in the western Kenyan highlands: a case-control study. Trop Med Int Health 14(10):1258–1265
15. Kazembe LN, Mathanga DP (2016) Estimating risk factors of urban malaria in Blantyre, Malawi: a spatial regression analysis. Asian Pac J Trop Biomed 6(5):376–381
16. Alem M, Enawgaw B, Gelaw A, Kena T, Seid M, Olkeba Y (2013) Prevalence of anemia and associated risk factors among pregnant women attending antenatal care in Azezo Health Center Gondar town, Northwest Ethiopia. J Interdiscip Histopathol 1(3):137–144
17. Subramaniyan K, George M, Seshadri D, Jena A, Chandraprabha. N (2016) Prevalence of anemia among health science students of a university in South India. Int J Res Med Sci 4(10):4598–4601
18. Alem M, Kena T, Baye N, Ahmed R, Tilahun S (2013) Prevalence of anemia and associated risk factors among adult HIV patients at the anti-retroviral therapy clinic at the University of Gondar Hospital, Gondar, Northwest Ethiopia. Sci Rep 2:662. https://doi.org/10.4172/scientificreports.662
19. Durge PM, Kasturwar NB (2012) Prevalence of anemia among adolescent girls in an urban slum. Natl J Commun Med 3(1):108–111
20. Gogoi I, Mahanta TG, Sarma R, Gogoi PP, Saikia H (2016) Prevalence and socio-demographic factors affecting anemia in pregnant women of Dibrugarh District, Assam, India. Indian J Commun Health 28(2):202–207