Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Weixing Huang Date

Factors Associated with Anemia among Adolescents

By

Weixing Huang

Master of Science in Public Health

Biostatistics and Bioinformatics

Renee H. Moore, PhD

Committee Chair

Melissa F. Young, PhD

Committee Member

Factors Associated with Anemia among Adolescents

By

Weixing Huang

B.E.
China Agricultural University
2016

Thesis Committee Chair: Renee H. Moore, PhD

Thesis Committee Member: Melissa F. Young, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics and Bioinformatics
2019

Abstract

Factors Associated with Anemia among Adolescents

By Weixing Huang

Background: Anemia is a condition in which the hemoglobin concentration in the blood is lower than normal; anemia affects nearly one third of the world's population. Anemia in adolescents can lead to impaired physical growth and mental development, resistance to infection, and reduced school performance and work capacity. So, it is important to assess the association between anemia and risk factors for anemia in adolescents.

Methods and Materials: Adolescents (age range: 10 to 19 years) who measured hemoglobin from 16 nationally representative cross-sectional surveys were analyzed by each country and pooled by the infection burden and risk in the country (n=20719). From these surveys, the prevalence of anemia was reported and univariate associations between anemia and factors including micronutrient deficiency, inflammation, malaria and demographic factors at every country level and by infection burden as well were examined. Univariate and multivariable logistic regression models were fit to identify key determinants of anemia in adolescents stratified by infection burden group.

Results: There was highly significant (P-value<0.0001) association between iron deficiency and anemia among adolescents from most country surveys excluding Mexico (2012) (P-value=0.74). This association was also highly significant among low, moderate, high country infection burden group (P-value <0.0001). In the multivariable analysis, anemia among adolescents who had iron deficiency (OR=6.08, p <0.001), any inflammation (OR=1.88, p-value=0.013), vitamin A-deficiency (OR=13.84, P=0.017), lower socioeconomic status (SES), (OR=2.07, p-value=0.013), lower education (OR=0.30, p-value<0.001) were associated with anemia in low infection burden group. Folate deficiency (OR=2.59, p-value<0.001), iron deficiency (OR=7.20, p-value<0.001) and any inflammation (OR=1.91, p-value<0.001) and with the increasing of age in one year (OR=1.20, p-value<0.001) were associated with anemia in moderate infection group. Folate deficiency (OR=2.48, p-value<0.001), Iron deficiency (OR=2.45, P-value=0.006), any inflammation (OR=2.14, P-value=0.012) were associated with anemia in high infection group.

Conclusion: Risk factors associated with anemia among adolescents vary according to a country's infection burden. In the multivariable analysis, iron deficiency and inflammation were consistently associated with anemia in low, moderate and high infection burden group. In order to improve anemia prevalence for adolescents, should consider assess both micronutrient deficiencies in different infection burden of the population.

Factors Associated with Anemia among Adolescents

By

Weixing Huang

B.E.
China Agricultural University
2016

Thesis Committee Chair: Renee H. Moore, PhD

Thesis Committee Member: Melissa F. Young, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics and Bioinformatics
2019

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude to my advisor Dr. Renee H. Moore, the director of Biostatistics Collaboration Core, Rollins School of Public Health. Throughout my thesis work, she has been my mentor and advisor, I am so lucky to have opportunity to do thesis under her guidance and enjoy every meeting with her to explore something new to improve my work.

I would also like to thank my thesis reader Dr. Melissa Young, the assistant professor of Department of Global Health at Emory University. From choosing topic of thesis to final revision, she paid every attention to keep steering me in the right direction and provided lots of constructive suggestions to improve thesis. Without her advisement, I would not able to finish this thesis.

I am so grateful to our Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) Team. Especially to Dr. Yaw Addo, Dr. Parminder S. Suchdev, Dr. Anne Williams, Emma Yu and Craig Ou. They gave me detailed revision advice to make my thesis more reasonable.

Finally, I must express my profound gratitude to my mother for supporting and encouraging me throughout my years of study. Her endless love and unfailing support have been great motivation for me all the time. Thank you and I love you!

CONTENTS

1. Introduction

Anemia is a condition in which the hemoglobin (Hb) concentration in the blood is lower than normal and it affects nearly one third of the world's population[1,2]. Anemia adversely affects people's cognition system and motor development along with low productivity and fatigue [3,4]Anemia is associated with a person's age, sex, pregnancy and altitude[5]. It is more common in developing countries, where adolescents are at a higher risk for the condition, constituting a serious public health problem [6]. Although there are certainly a lot of studies on risk factors of the anemia, they more focused more on population of women of reproductive age [7] or pre-school aged children [8], rather than adolescents. The prevalence of anemia among adolescents is 15% worldwide, 27% in developing countries and 6% in developed countries[9]. Causes of anemia in developing countries are multi-factorial, which include nutritional (iron, folate, vitamin A and vitamin B12) deficiencies, infections such as malaria, and chronic illness [10]. The prevalence of iron deficiency and subsequent anemia increases from the start of adolescence. As for girls, it is caused by increased requirements of nutrition for growth and exacerbated several years later by the start of menstruation, but not for boys. The reason why anemia increases from the start of adolescents is also due to rapid growth with increase blood volume, lean body mass and red cell mass, which increase the requirements of iron for myoglobin in their muscles and Hb [11]. Therefore, the physical and physiological changes which occur during adolescents make more requirements on nutrition and make them more vulnerable to nutritional deficiencies. The level of iron requirements increases from the level of 0.7-0.9 mg

iron/day to as much as 1.40-3.27 mg iron/day in adolescent girls and 1.37-1.88 mg iron/day in adolescent boys[12].

Anemia in adolescents can cause impaired physical growth, resistance to infection and mental development and reduced school performance and work capacity[13]. The major cause of anemia in most populations is iron deficiency and it has been long considered that iron deficiency contributes up to 50% of all anemia[6]. Also, a recent study which used meta-analysis on the response of hemoglobin to iron-fortification indicated that iron deficiency was the major contributor to anemia among multiple geographic settings[14]. The objective of this study is to assess the association between anemia and possible risk factors for anemia among adolescents of boys and girls in 19 country surveys.

2. Method

The data for analysis are from the BRINDA project (www.BRINDA-nutrition.org), which included surveys that were conducted after 2004 and included the measurement of anemia (hemoglobin), biomarker of iron (based on ferritin or soluble transferrin receptor (sTfR) or vitamin A status (retinol-binding protein (RBP) or retinol) and inflammation (α-1-acid glycoprotein (AGP) or C-reactive protein (CRP)). The protocol was reviewed by the institutional review boards of the NIH [15] and Emory University Institutional Review Board (IRB). This analysis includes surveys with measures of hemoglobin in adolescents (age range: girls from 10-19y, boys from 10-15y), and examined data from 19 nationally representative cross-sectional surveys from 5 countries in the region of America (Colombia, Mexico, the United States, Ecuador, Malawi), 3 countries in West Africa (Cameroon, Côte d'Ivoire, Liberia), 2 countries in Southeast Asia (Bangladesh, Laos), 3 countries in European Region (Georgia, Azerbaijan, the United Kingdom), 2 countries in Eastern Mediterranean Region (Pakistan, Afghanistan), and 3 countries in Western Pacific Region (Cambodia, Vietnam, Papua New Guinea). Additionally, participants from Azerbaijan, Cote d'Ivoire, Georgia, Laos, Liberia, PNG and Vietnam are all female. Cambodia, Afghanistan, Pakistan and Cameroon were excluded from analysis since the sample size for adolescent was too small (n< 100) to conduct univariate analysis. Original surveys' descriptions and relevant references have been previously reported[16].

2.1 Grouping Countries by Infection Burden

Since causes of anemia may vary depending on environmental and socioeconomic characteristics and the intensity of exposure to infections and to inflammation-inducing conditions. The surveys were categorized into 3 groups (low, moderate, high) representing risk and burden of infectious disease and inflammation (hereinafter referred to as infection burden); the data analysis was conducted separately for each infection burden group. Countries were assigned to infection groups by adapting the approach developed by Petry et al[17]. and applied by Wirth et al. [7] and Engle-Stone et al. [8], whereby national-level prevalence estimates of malaria, HIV infection, access to improved drinking water and sanitation facilities, and schistosomiasis were used to calculate an equally weighted infection score for each country and to group countries on the basis of their infection scores (see **Supplemental Table 1** for details). Since obesity prevalence was not reported for many countries, our determination of infection burden group differed slightly from the method used by Petry et al.[17] . For our study, Côte d'Ivoire, Liberia, Laos, and Papua New Guinea, Malawi were classified as countries with a high infection burden; Colombia, Mexico, Ecuador, Vietnam, Azerbaijan, Bangladesh, Pakistan were classified as countries with a moderate infection burden; and Georgia, United Kingdom and the United States were classified as having a low infection burden.

2.2 Determinants of Anemia and Case Definitions

The WHO definitions of anemia status were used to classify anemia in adolescents [5]. Any anemia was defined as a hemoglobin concentration <115 g/L for adolescents whose age was under 12, hemoglobin concentration <120 g/L for adolescents whose age is between 12-14 and girls greater than 15 years old, hemoglobin concentration <130 g/L for boys greater than 15 years old. Severe anemia was defined as hemoglobin concentrations <80

g/L for adolescents. For girls among 15 years old, hemoglobin concentrations were adjusted for altitude and the intensity of cigarette smoking according to WHO procedures [5] in the Colombia, Georgia, Papua New Guinea, and Mexico 2006, and Mexico 2012 surveys. In Laos, hemoglobin was adjusted only for altitude and in United States, hemoglobin was adjusted only for smoking. No adjustment to hemoglobin were made for data from Côte d'Ivoire, Cameroon and Liberia. Hemoglobin concentrations were considered biologically implausible when values were out of range (40–180 g/L); these values were set to missing. Utilizing the WHO classification of the public health significance of anemia [6], an anemia prevalence of <5% was considered normal, prevalence of 5.0–19.9% was used to denote a mild public health problem, a prevalence of 20.0–39.9% was used to denote a moderate problem, and a prevalence of ≥40% was used to denote a severe problem.

Ferritin was used as the indicator of iron status in population-based surveys according to WHO [18.] Ferritin was adjusted for inflammation as measured by AGP and CRP concentrations with the use of the internal regression correction (IRC) approaches that are defined by Namaste et al.[19] . Ferritin was adjusted with the use of both CRP and AGP when available, and with only CRP in countries for which AGP was not reported. The advantage of IRC approach is it can provide a continuous adjustment of ferritin concentrations and result in a greater difference from the unadjusted prevalence as previous methods that adjusted ferritin when inflammation levels reached certain thresholds[20,21] . RBP or retinol were used as measures of vitamin A status. Though RBP and retinol concentrations are affected by inflammation, the findings by Larson et al. [22].showed that the relations between these indicators and inflammatory markers in

BRINDA data set was inconsistent so in this study we did not adjust RBP or retinol concentrations for inflammation.

Iron deficiency was defined as an adjusted ferritin concentration <15 μg/L [18], Iron-deficiency anemia was defined as concurrent iron deficiency (ferritin concentration <15 μg/L) and anemia. Vitamin A deficiency was defined as RBP or retinol concentrations <0.7 μmol/L[23], folate deficiency was defined as a plasma or serum folate concentration <10 nmol/L, and vitamin B-12 deficiency was defined as a serum cobalamin concentration <150 pmol/L [24]. In this study, folate and vitamin B-12 concentrations were not corrected for inflammation.

CRP was measured in all surveys while AGP was measured in some of the country surveys (Azerbaijan, Bangladesh, Cote d'Ivoire, Laos, Liberia, Malawi and Papua New Guinea). The inflammation status was classified into 2 categories as follows: no inflammation and any inflammation. No inflammation was defined as having both a normal CRP concentration (≤5 mg/L) and a normal AGP concentration (≤1.0 g/L). When there are no AGP values, no inflammation was defined as having a normal CRP concentration. Survey subjects with no biochemical measure of inflammation were recoded as missing values. Accordingly, ferritin values were also recoded as missing if no data on inflammation was available.

Three household-level factors which were related to water, sanitation and socioeconomic status (SES) were examined as potential risk factors of anemia status. Water and sanitation indicators were defined according to UNICEF-WHO [25] guidelines by classifying the household drinking water source and household sanitation facility as being

improved or unimproved. Household SES was being classified as low SES, medium SES and high SES, which was derived from asset scores from original surveys. Because of a lack of data on household assets or income, SES could not be calculated in Georgia and Vietnam.

2.3 Statistical Method

2.3.1 Analysis by Country Survey Level

2.3.1.1 Descriptive Analysis

All analyses were conducted with SAS 9.4 (SAS Institute). For continuous variables (hemoglobin concentration, age in years), the mean, minimum and maximum were reported. For binary and categorical variables, prevalence of anemia and severe anemia, prevalence of micronutrient deficiencies (iron deficiency, iron-deficiency anemia, vitamin A deficiency, folate deficiency, vitamin B-12 deficiency) inflammation and malaria, the frequencies and percentage (95 % Wald confidence interval) were presented.

2.3.1.2 Univariate Association Analysis

Univariate analyses of the associations between anemia and demographic and nutrition factors were conducted for each country survey. The prevalence of anemia was calculated for each nutrient risk factors (iron deficiency, Vitamin A deficiency, folate deficiency, Vitamin B-12 deficiency, malaria, any inflammation) and sociodemographic factors (residence type, socioeconomic status, sanitation facility and water source). The significance of the difference between anemia prevalence for each factor's subgroups was examined utilizing Pearson's chi-square test and when 20% or more of expected cell

sizes <5, Fisher's exact test was utilized. An alpha level of 0.05 was used to determine statistical significance.

2.3.2 Analysis by Pooled Data

2.3.2.1 Univariate and Multivariate Logistic Regression Analysis

Since there were some missing values with each micronutrient biomarkers, the first approach to deal with pooled data to split data based on whether or not they have each important biomarker. As a result, we conducted univariate logistic regression of factors related to anemia who has RBP, ferritin, serum folate, vitamin B-12, inflammation status and malaria status for each infection burden group respectively. The model is shown here:

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_i x_i,$$

Where i =age, sex, iron deficiency, vitamin A deficiency, folate deficiency, vitamin B-12 deficiency, any inflammation, malaria, residency type, socioeconomic status, sanitation facility, water source, education level and π indicates the probability of an event (here it is the prevalence of anemia).

Factors with associations with the anemia at alpha level equal to 0.15 in the univariate analyses for each infection burden country group were considered for the multivariate logistic regression models.

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m$$

Where 1,2…m indicates covariates put in the multivariable regression model and π indicates the probability of an event (here it is the prevalence of anemia).

We also constructed multivariable logistic model for each micronutrient biomarker and health condition (Vitamin A, ferritin, folate, vitamin B-12, inflammation, malaria). The models contained main effect (each micronutrient biomarker and demographic factors, environmental factors) and interaction term between main effect and infection burden group (low, moderate, high). First, we removed non-significant ($p > 0.05$) interaction term from the multivariable logistic regression model 1. Then we re-fitted the multivariable logistic regression model without non-significant interactions and then removed non-significant ($p>0.05$) main effects from the model to get the final model with all main effects and interactions included at alpha level equal to 0.05.

Second approach to deal with pooled data is applying 80% rule. After fitting with univariate logistic regression for each covariate, we removed covariates which were non-significant (SES and water source, p-value >0.05) in the univariate model and with missing data above 20%. Then we built multivariable logistic regression model with remaining covariates (Iron deficiency, any inflammation, sex, residence type, age and infection burden group). Main effect of sex was non-significant (P-value >0.05) in the multivariable model so we decided to take out variable sex (P-value>0.05) and re-built multivariable logistic model. Finally, there were 5 variables (Iron deficiency, any inflammation, residence type, age and infection burden group) remaining in the model. In order to consider interaction between infection burden group and other covariates, we constructed multivariable model with 5 main effects and interaction term between

infection burden group and other 4 covariates. But since there was the issue of multicollinearity (VIF >10), type 3 analysis didn't work well. So we decided to build model with each covariates (age, sex, iron deficiency, vitamin A deficiency, folate deficiency, vitamin B-12 deficiency, any inflammation, malaria, residency type, socioeconomic status, sanitation facility, water source, education) and infection burden to see whether there was interaction between main effects and infection burden groups (**Supplement table 2**).

2.3.3 Analysis by Infection Burden Group

2.3.3.1 Descriptive Analysis

For continuous variables (hemoglobin concentration, age in years), the mean, minimum and maximum were reported. For binary and categorical variables, prevalence of anemia, prevalence of micronutrient deficiencies (iron deficiency, iron-deficiency anemia, vitamin A deficiency, folate deficiency, vitamin B-12 deficiency) inflammation and malaria, the frequencies and percentage (95 % Wald confidence interval) were presented.

2.3.3.2 Univariate Association Analysis

Univariate analyses of the associations between anemia and demographic and nutrition factors were conducted for each infection burden group. The prevalence of anemia was calculated for each nutrient risk factors (iron deficiency, Vitamin A deficiency, folate deficiency, Vitamin B-12 deficiency, malaria, any inflammation) and sociodemographic factors (residence type, socioeconomic status, sanitation facility and water source). The significance of the difference between anemia prevalence for each factor's subgroups was

examined utilizing Pearson's chi-square test and when 20% or more of expected cell sizes <5, Fisher's exact test was utilized. An alpha level of 0.05 was used to determine statistical significance.

2.3.3.3 Univariate and Multivariate Logistic Regression Analysis

Because there was interaction between infection burden and other main effects, we constructed univariate logistic model stratified by infection burden (Low, moderate, high). Then built multivariable model for each infection burden group, with removing covariates which were non-significant (P-value >0.05) in the univariate model. After taking out non-significant covariates in the multivariable model, we got the final model for low, moderate, high infection group respectively.

3. Results

3.1 Result of Analysis by Country Level

3.1.1 Result of Descriptive Statistics

In our study, the sample was restricted to adolescents (age between 10 to 19) with hemoglobin concentration values reported. The descriptive statistics for participants' characteristics are reported in **Table 1.** This analysis includes 20719 adolescents who had valid age and hemoglobin measurements. At the country level, the mean of age ranged from 11.03 to 17.51 y. The prevalence of anemia ranged from 5.14% (in UK) to 59.09% (in Cote d'Ivoire), at the country level.

The prevalence of micronutrient deficiency, inflammation and malaria were shown as **Table 2a-c.** The estimated prevalence of iron deficiency on the basis of inflammation-adjusted ferritin values ranged from 0% (Georgia) to 38.62% (Liberia) in all countries except for PNG which didn't have serum ferritin value to define iron status. As for the prevalence of iron-deficiency anemia, 4 countries (Azerbaijan, Cote d'Ivoire, Laos and Liberia) are >10%.

Among 11 country surveys measured RBP or retinol in adolescents. In Bangladesh (18.50%), Ecuador (10.87%), Malawi (4.52%) and Vietnam (2.17%), the prevalence of vitamin A deficiency is >2 % (Table 2b). The prevalence of any inflammation ranged from 4.7% in Vietnam to 34.55% in Cote d'Ivoire. Only 3 countries, Cote d'Ivoire, Liberia and Malawi had the malaria status. The prevalence of malaria is similar in Malawi and Liberia but much lower in Cote d'Ivoire where only current malaria was assessed with the use of microscopy. Only a few country surveys measured serum or plasma folate. While the prevalence of folate deficiency is high in Côte d'Ivoire (87.74%) and Georgia (76.92%), it is lower (<3%) in Ecuador, Mexico (2012) and US. Similarly, the prevalence of vitamin B-12 was very low (< 3%) in Ecuador, Mexico (2012) and US.

3.1.2 Result of Univariate Association Analysis

Table 3a reports the results of the Pearson's Chi square or Fisher's exact tests for the association with iron deficiency and anemia. The null hypothesis is H_0: There is no association between iron deficiency and anemia. By computing corresponding p-value with the significance level a=0.05, we found that for every country, except for Mexico (2012), anemia was significantly associated with the iron deficiency. In other words, the

prevalence of anemia was significantly higher in iron-deficient adolescents in all countries except for Mexico (2012). Vitamin-A deficiency was associated with prevalence of anemia in Bangladesh and US, but not in Azerbaijan, UK, Malawi, Mexico (2012) and Vietnam. The prevalence of anemia was only significantly associated with folate deficiency in Azerbaijan, Bangladesh and Vietnam **(Table 3b).**

Vitamin B-12 status was also associated with the prevalence of anemia in some of the countries (Ecuador, Colombia, Malawi, US). (**Table 3b.)** Mixed results were observed for any inflammation as well **(Table 3c)**, there were significant ($P<0.05$) association between any inflammation and anemia in some of the country surveys (Cote d'Ivoire, Colombia, Ecuador, Liberia, Mexico (2006), Mexico (2012) and US).

The demographic factors that were reported varied by country survey as **Table 3d-f.** Only some of the countries had an association with the prevalence of anemia and demographic factors and it was not consistent. Water source status (unimproved or improved) was associated with the prevalence of anemia only in Cote d'Ivoire (P-value =0.02)

3.2 Results of analysis by pooled data

3.2.1 Result of Univariate and Multivariate Logistic Regression Analysis

The first approach we mentioned in the method part was not used since sample size was too small for each model and the result of each model was not representative of our original dataset. Therefore, we only focused on second approach which applied 80% rule.

In **Table 4**, we presented the univariate associations of anemia with each covariate, including vitamin A, iron, folate, vitamin B-12, inflammation, malaria, sex, SES, residence type, sanitation source, water source, infection burden group, education, age. We set levels which are less likely to have anemia as reference. The result showed that except for SES, water source and malaria, all other covariates had significant association with anemia (P-value < 0.05).

When we saw if there is any interaction between infection burden and each covariate (age, sex, iron deficiency, vitamin A deficiency, folate deficiency, vitamin B-12 deficiency, any inflammation, malaria, residency type, socioeconomic status, sanitation facility, water source, education), 7 of 12 sub-models had significant (P-value <0.05) interaction with infection burden (Supplement table 2)

3.3 Result of analysis by infection burden group

3.3.1 Result of Descriptive Statistics

It is noticeable that in high-infection countries the prevalence of anemia was much higher (33.72%) than in moderate (7.31%) or lower countries (6.50%). The prevalence of severe anemia (hemoglobin concentrations <80 g/L) in all of the countries were lower than 2%, so we decided to not consider it in later analysis.

3.3.2 Result of Univariate Association Analysis

When it comes to infection burden group, prevalence of anemia was significantly associated with iron deficiency in low, moderate and high groups (P-value < 0.0001). There were significant associations between anemia and Vitamin-A deficiency in low (P-

value=0.050) and moderate countries (P-value=0.0001) but not high infection burden. (**Table 3a.**) There was strong association (P-value <0.0001) in moderate and high infection burden groups. (**Table 3b.)** There were also strong association (P-value<0.0001) with the prevalence of anemia and any inflammation when grouping countries by infection burden. (**Table 3c.)** Though the prevalence of anemia was consistently higher in adolescents with malaria, these differences were not significant in Cote d'Ivoire and Liberia and in the pooled analysis. (**Table 3c.)**

SES showed a significant association with anemia prevalence in all infection burden groups. (**Table 3f.)** Water source status (unimproved or improved) was not associated with the prevalence of anemia in all infection burden groups. (**Table 3e.)** We also found that there was higher prevalence of iron deficiency in high-infection countries than moderate or lower infection countries.

3.3.3 Result of Univariate and Multivariate Logistic Regression Analysis

Results from the univariate analyses for each infection burden group are presented in **Table 5, 6, 7.** In the low-infection group, vitamin A, iron, vitamin B-12, inflammation, sex, SES, education level and age showed significant effect on anemia prevalence (P-value <0.05) and in moderate-infection group, vitamin A, iron, folate, vitamin B-12, inflammation, sex and age, showed significant effect (P-value <0.05). As for high-infection group, iron, folate, inflammation, sex, SES, toilet source and age showed significant effect (P-value <0.05).

Results from multivariable analyses for each infection groups are presented in **Table 8, 9, 10.** In the low infection group, the odds of anemia in iron-deficient adolescents was 6.08

times higher than that of iron-replete ones (P-value<0.001, OR 95% CI: 4.26-8.68). Adolescents with any inflammation had 1.88 times the odds of being anemic than that of adolescents with no inflammation (P-value=0.01, 95% OR CI: 1.16-3.03). And adolescents with low SES had 2.07 times of odds of being anemic compared to high SES group (P-value=0.013, 95% OR CI: 1.17-3.68). Adolescents who was in deficiency of vitamin A had 13.84 times odds of being anemic than sufficient group (P-value=0.017, OR 95% CI: 1.58- 120.91).

As for moderate infection group, only folate deficiency, iron deficiency, any inflammation is associated with prevalence of anemia in multivariable analysis (P-value <0.001). The odds of anemic in adolescents with iron deficiency had 7.20 than in iron sufficient group (P-value <0.001, OR 95% CI: 1.87-3.59). Two important risk factors, iron deficiency and any inflammation increased odds of anemic in adolescents in 7.20 (OR 95% CI: 5.81-8.94, P-value <0.001) times and 1.91 (OR 95% CI: 1.40-2.60) times compared to iron-sufficient and no inflammation group. With the increasing of age for one year, there is 1.20 times the odds to be diagnosed with anemia among adolescents.

Folate deficiency, iron deficiency and any inflammation remained significant in the final multivariable logistic model for high infection burden group (P-value <0.05). Adolescents who were in folate deficiency had 2.48 (OR 95% CI: 1.46-4.18, P-value <0.001) times the odds of being anemic than reference. Iron-deficient and with any inflammation adolescents had 2.45 (OR 95% CI: 1.29-4.64, P-value=0.006) and 2.14 (OR 95% CI: 1.18-3.87, P-value=0.012) times of odds being anemic than adolescents who are in iron sufficient and without inflammation.

4. Conclusion

According to our study, the contribution of iron deficiency to anemia varied based on a country's infection burden. However, the association between iron deficiency and anemia were all highly significant among low, moderate, high country infection burden group. Per the multivariable logistic regression models, iron deficiency and any inflammation had consistently significant associations with anemia among adolescents. Iron deficiency and any inflammation were the only robust finding; there were different predictors for each infection burden group. For the low infection burden group, vitamin A deficiency, low socioeconomic status and high education had significant association with anemia. In the moderate infection group, ferritin deficiency and older age had significant association with anemia and in high infection group, ferritin deficiency had significant association with anemia.

5. Discussion

The study provides evidence for associations between prevalence of anemia and vitamin A deficiency, folate deficiency, iron deficiency, inflammation, socioeconomic status, education and age. Although we did find an association between vitamin A status and anemia in the low infection group which including Georgia, UK and the United States (Vitamin A status was not available in Georgia) the 95% CI was extremely wide: OR=13.84 (with 95% CI:1.58, 120.91). The reason might be we had insufficient data for vitamin A-deficient group. In UK, only 4 of 506 adolescents were in vitamin A deficiency and in the US, there was only 11 of 3152 adolescents were deficient in vitamin A. Despite of small p-value (P-value=0.017), such unbalanced data in vitamin A status

resulted in crude and imprecise estimate of odds ratio for vitamin A covariate in low infection burden group.

In the low infection burden group, education status was considered as covariate in the final model while it was removed from moderate and high infection burden group models. It was noticeable that adolescents with lower education had lower odds (OR=0.30, (95% CI: 0.14-0.62)) of being anemic. But information of education status was only available in the US in the low infection burden group, so it was not as much as representative. Moreover, since our respondent were adolescents, who were undertaking education in the meantime. In this case, considering education as factors with anemia might be the same as interpretation of age. Therefore, it seemed education status was not reasonable considered to be a covariate in this study.

Although the interpretation of ferritin concentrations is complicated by the presence of inflammation, our analysis took advantage of newer methods for mathematically adjusting these indicators for inflammation, namely, the use of a regression approach rather than the application of correction factors on the basis of infection categories as proposed previously by Thurnham et al. [20,21.] The regression approach resulted in a greater reduction in the prevalence of iron deficiency than shown with correction-factor approaches; this difference in iron assessment may explain some inconsistencies with the results of other studies.

Handling with missing data and construct an interpretable multivariable model had been major issue for this study. Although our sample size was 20719 in total, there was no observation who had every covariate in the meantime, so there was no way to construct

pool multivariable model for the whole population, which indicated we need to construct stratified models. Then we tried to assess interaction term before stratifying, but there was strong multicollinearity (VIF>100) among infection burden group (low, moderate and high) with iron status (deficiency and sufficiency), inflammation (any and no) and other covariates. Finally, the way to address this issue, we looked into every sub-model which only contain every main effect and interaction with infection burden group (Supplement table2). There were 7 among 12 covariates showed significant interaction with infection burden group, which supported stratify pooled population based on infection burden group.

Our analysis has some notable limitations. Firstly, data were cross-sectional, which prevented any temporal analysis of causation. Secondly, it will be better to apply complex survey design for analysis, but we didn't apply any strata and weight in this study. Because adolescent data were re-constructed sample from women of reproductive age (15-49y) and school aged children (5-15y), so the sampling frame such as weight, strata and cluster were not following with our target population-adolescents. We just got the prevalence of anemia and estimated it with 95% confidence interval by country level, and the survey's completion date were varying, which loses representative of sample. Thirdly, not all data sets contained measure of AGP, which is one of most important biomarkers in the study of anemia and comprised the measurement of inflammation.

6. Reference

1. FAO. The state of food security and nutrition in the world. *FAO, IFAD WFP*. 2017. doi:I4646E/1/05.15
2. Kassebaum NJ, Fleming TD, Flaxman A, et al. The Global Burden of Anemia. *Hematol Oncol Clin North Am*. 2016. doi:10.1016/j.hoc.2015.11.002
3. Balarajan Y, Ramakrishnan U, Özaltin E, Shankar AH, Subramanian S V. Anaemia in low-income and middle-income countries. *Lancet*. 2011. doi:10.1016/S0140-6736(10)62304-5
4. Haas JD, Brownlie, Thomas I. Iron Deficiency and Reduced Work Capacity: A Critical Review of the Research to Determine a Causal Relationship. *J Nutr*. 2001.
5. World Health Organization. *Haemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity*.; 2011. doi:2011
6. (WHO) WHO, (UNICEF) TUNCF, (UNU) ATUNU. Iron deficiency anaemia: Assessment, Prevention, and Control: A guide for programme managers. *World Heal Organ*. 2001. doi:10.1136/pgmj.2009.089987
7. Wirth JP, Woodruff BA, Engle-Stone R, et al. Predictors of anemia in women of reproductive age: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr*. 2017. doi:10.3945/ajcn.116.143073
8. Engle-Stone R, Aaron GJ, Huang J, et al. Predictors of anemia in preschool children: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr*. 2017. doi:10.3945/ajcn.116.142323
9. Dugdale M. Anemia. *Obstetrics and Gynecology Clinics of North America* 2001;28(2):363-81

10. Stevens GA, Finucane MM, De-Regil LM, et al. Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995-2011: A systematic analysis of population-representative data. *Lancet Glob Heal*. 2013. doi:10.1016/S2214-109X(13)70001-9

11. Usha R. Nutritional Anemia. *Boca Raton, FL: CRC press*; 2001: 8–12.

12. World Health Organization. Prevention of iron deficiency anaemia in adolescents. *Searo*. 2011. doi:10.1109/VTC.1982.1623054

13. Sandler AD. Iron deficiency and cognitive achievement among school-aged children and adolescents in the United States. *J Dev Behav Pediatr*. 2001. doi:10.1097/00004703-200112000-00028

14. Kassebaum NJ, Jasrasaria R, Naghavi M, et al. A systematic analysis of global anemia burden from 1990 to 2010. *Blood*. 2014. doi:10.1182/blood-2013-06-508325

15. Suchdev PS, Namaste SM, Aaron GJ, Raiten DJ, Brown KH, Flores-Ayala R. Overview of the Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) Project. *Adv Nutr*. 2016. doi:10.3945/an.115.010215

16. Namaste SM, Aaron GJ, Varadhan R, Peerson JM, Suchdev PS. Methodologic approach for the Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr*. 2017. doi:10.3945/ajcn.116.142273

17. Petry N, Olofin I, Hurrell RF, et al. The proportion of anemia associated with iron

deficiency in low, medium, and high human development index countries: A systematic analysis of national surveys. *Nutrients*. 2016. doi:10.3390/nu8110693

18. WHO. Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations. Vitamin and Mineral Nutrition Information System. *Inside,VMNIS*. 2011. doi:(WHO/NMH/NHD/MNM/11.2)
19. Namaste SM, Rohner F, Huang J, et al. Adjusting ferritin concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr*. 2017. doi:10.3945/ajcn.116.141762
20. Thurnham DI, McCabe LD, Haldar S, Wieringa FT, Northrop-Clewes CA, McCabe GP. Adjusting plasma ferritin concentrations to remove the effects of subclinical inflammation in the assessment of iron deficiency: A meta-analysis. *Am J Clin Nutr*. 2010. doi:10.3945/ajcn.2010.29284
21. Thurnham DI, McCabe GP, Northrop-Clewes CA, Nestel P. Effects of subclinical infection on plasma retinol concentrations and assessment of prevalence of vitamin a deficiency: Meta-analysis. *Lancet*. 2003. doi:10.1016/S0140-6736(03)15099-4
22. Larson LM, Namaste SM, Williams AM, et al. Adjusting retinol-binding protein concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr*. 2017. doi:10.3945/ajcn.116.142166
23. Tanumihardjo S.A. Biomarkers of vitamin A status: What do they mean? Panama City, Panama, 15–17 September 2010. World Health Organization: Geneva, Switzerland.;2012

24. WHO, FAO. *Guidelines on Food Fortification with Micronutrients*.; 2006. doi:10.1242/jeb.02490

25. WHO/UNICEF. *2015 ANNUAL REPORT WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation*.; 2015. doi:10.1007/s13398-014-0173-7.2

7. Tables

Table 1 Age, sex and anemia prevalence in adolescents by country and infection burden[1]

Country (Year)	Sample size	Age in years Mean (Min, max)	Female N	Female %	Anemia N	Anemia 95% CI
Azerbaijan (2013)	361	17.51 (15.00, 19.92)	361	100	127	35.18 (30.25, 40.11)
Bangladesh (2012)	794	12.67 (10.00, 19.00)	487	61.34	126	15.87 (13.33, 18.41)
Cote d'Ivoire (2007)	110	17.31 (15.00, 19.00)	110	100	65	59.09 (49.90, 68.28)
Colombia (2010)	6953	14.11 (10.00, 19.92)	5539	79.66	372	5.35 (4.82, 5.88)
Ecuador (2012)	4151	13.81 (10.00, 19.00)	2718	65.48	228	5.49 (4.80, 6.19)
UK (GB2014)	545	17.05 (15.00, 19.00)	364	66.79	28	5.14 (3.28, 6.99)
Georgia (2009)	178	13.86 (10.00, 19.00)	178	100	45	25.28 (18.90, 31.67)
Laos (2006)	170	16.68 (15.00, 19.00)	170	100	69	40.59 (33.21, 47.97)
Liberia (2011)	378	17.83 (15.00, 19.92)	378	100	149	39.42 (34.49, 44.34)
Malawi (2016)	509	13.71 (10.00, 19.00)	327	64.24	101	19.84 (16.38, 23.31)
Mexico (2006)	1891	13.97 (10.00, 19.98)	1567	82.87	190	10.05 (8.69, 11.40)
Mexico (2012)	1110	11.03 (10.00, 19.53)	531	47.84	74	6.67 (5.20, 8.13)
PNG (2005)	132	16.89 (15.00, 19.00)	132	100	54	40.91 (32.52, 49.30)
US (2006)	3246	14.41 (10.00, 19.92)	2214	68.21	185	5.70 (4.90, 6.50)
Vietnam (2010)	191	16.84 (15.00, 19.00)	191	100	13	6.81 (3.23, 10.38)
Infection burden[2]						
Low	3969	14.45 (10.00, 19.92)	2756	69.44	258	6.50 (5.73, 7.27)
Moderate	15451	13.83 (10.00, 19.98)	11394	73.74	1130	7.31 (6.90, 7.73)
High	1299	15.92 (10.00, 19.92)	1117	85.99	438	33.72 (31.15, 36.29)

1. Any anemia was defined as a hemoglobin concentration <115 g/L for adolescents whose age is under 12, hemoglobin concentration <120 g/L for adolescents whose age is between 12-14 and girls greater than 15 years old, hemoglobin concentration <130 g/L for boys greater than 15 years old.
2. Countries were categorized by infection burden as follows—low: Georgia, United Kingdom and the United States; moderate: Colombia, Mexico (2006 and 2012), Ecuador, Vietnam, Azerbaijan, Bangladesh; high: Côte d'Ivoire, Liberia, Laos, and Papua New Guinea, Malawi.

Table 2 a: Prevalence of iron deficiency and iron-deficiency anemia in adolescents by country [1]

Country(year)	Iron deficiency		Iron-deficiency anemia	
	n/N	% (95% CI)	n/N	% (95% CI)
Azerbaijan (2013)	134/361	37.12 (32.14, 42.10)	79/361	21.88 (17.62, 26.15)
Bangladesh (2012)	72/783	9.20 (7.17, 11.22)	27/792	3.41 (2.15, 4.67)
Cote d'Ivoire (2007)	18/110	16.35 (9.45, 23.28)	15/110	13.64 (7.22, 20.05)
Colombia (2010)	1173/6953	16.87 (15.99, 17.75)	135/6952	1.94 (1.62, 2.27)
Ecuador (2012)	374/4150	9.01 (8.14,9.88)	119/4151	2.87 (2.36, 3.37)
UK (GB2014)	109/525	20.76 (17.29, 24.23)	16/544	2.94 (1.52, 4.36)
Georgia (2009)	0/178	0	0/178	0
Laos (2006)	61/170	35.88 (28.67, 43.09)	32/170	18.82 (12.95, 24.70)
Liberia (2011)	146/378	38.62 (33.72, 43.53)	86/378	22.75 (18.53, 26.98)
Malawi (2016)	50/509	9.82 (7.24, 12.41)	21/509	4.13 (2.40, 5.85)
Mexico (2006)	476/1884	25.27 (23.33, 27.23)	76/1890	4.02 (3.14, 4.91)
Mexico (2012)	165/1106	14.92 (12.82, 17.02)	12/1110	1.08 (0.47, 1.69)
PNG (2005)	NA	NA	NA	NA
US (2006)	423/1878	22.52 (20.63, 24,41)	96/3222	2.98 (2.39, 3.57)
Vietnam (2010)	31/191	16.23 (11.00, 21.46)	10/191	5.24 (2.08, 8.39)
Infection burden				
Low	532/2581	20.61 (19.05, 22.17)	112/3944	2.84 (2.32, 3.36)
Moderate	2425/15428	15.72 (15.14, 16.29)	458/15448	2.96 (2.70, 3.23)
High	275/1167	23.56 (21.13, 26.00)	154/1245	12.3710.54, 14.20)

1. Values in parentheses are 95% CIs. Iron deficiency was defined as an inflammation-adjusted ferritin concentration <15 µg/L. Iron-deficiency anemia was defined as a hemoglobin concentration <120 g/L and an inflammation-adjusted ferritin concentration <15 µg/L, NA, not available.

Table 2 b: Prevalence of Vitamin A, Folate and Vitamin B-12 deficiencies in adolescents by country[1]

Country (year)	Vitamin A deficiency		Folate deficiency		Vitamin B-12 deficiency	
	n/N	% (95% CI)	n/N	% (95% CI)	n/N	% (95% CI)
Azerbaijan (2013)	4/361	1.11 (0.03, 2.19)	132/345	38.26 (33.13, 43.39)	38/173	21.97 (15.80, 28.13)
Bangladesh (2012)	146/789	18.50 (15.79, 21.21)	68/160	42.50 (34.84, 50.16)	4/164	2.44 (0.08, 4.80)
Cote d'Ivoire (2007)	0/110	0	93/106	87.74 (81.49, 93.98)	11/56	19.64 (9.24, 30.05)
Colombia (2010)	NA	NA	NA	NA	137/2965	4.62 (3.87, 5.37)
Ecuador (2012)	10/92	10.87 (0.45, 17.23)	30/4151	0.72 (0.46, 0.98)	39/4062	0.96 (0.66, 1.26)
UK(GB2014)	4/506	0.79 (0.02, 1.56)	NA	NA	20/516	3.88 (2.21, 5.54)
Georgia (2009)	NA	NA	20/26	76.92 (60.73, 93.12)	NA	NA
Laos (2006)	NA	NA	NA	NA	NA	NA
Liberia (2011)	5/378	1.32 (0.17, 2.47)	NA	NA	NA	NA
Malawi (2016)	23/509	4.52 (2.71, 6.32)	30/160	18.75 (12.70, 24.80)	18/160	11.25 (6.35, 16.15)
Mexico (2006)	NA	NA	NA	NA	NA	NA
Mexico (2012)	12/828	1.45 (0.64, 2.26)	2/1105	0.18 (0.00, 0.43)	7/1105	0.63 (0.17, 1.10)
PNG (2005)	0/132	0	NA	NA	NA	NA
US (2006)	11/3152	0.35 (0.14, 0.55)	26/3226	0.81 (0.50, 1.11)	21/3223	0.65 (0.37, 0.93)
Vietnam (2010)	4/184	2.17 (0.07, 4.28)	23/177	12.99 (8.04, 17.95)	3/52	5.77 (0.00, 12.11)
Infection burden						
Low	15/3658	0.41 (0.20, 0.62)	46/3252	1.41 (1.01, 1.82)	41/3739	1.10 (0.76, 1.43)
Moderate	176/2254	7.81 (6.70, 8.92)	255/5938	4.29 (3.78, 4.81)	228/8521	2.69 (2.33, 3.02)
High	28/1129	2.48 (1.57, 3.39)	123/266	46.24 (40.25, 52.23)	29/216	13.43 (8.88. 17.97)

1. Vitamin A deficiency was defined as a retinol-binding protein or retinol concentration < 0.7µmol/L. Folate deficiency was defined as a folate concentration <10 nmol/L. Vitamin B-12 deficiency was defined as a vitamin B-12 concentration <150 pmol/L NA, not available.

Table 2 c: Prevalence of inflammation and malaria in adolescents by country

Country(year)	Inflammation[1]		Malaria	
	n/N	% (95% CI)	n/N	% (95% CI)
Azerbaijan (2013)	74/361	20.50 (16.33, 24.66)	NA	NA
Bangladesh (2012)	95/794	11.96 (9.71, 14.22)	NA	NA
Cote d'Ivoire (2007)	38/110	34.55 (25.66, 43.43)	5/109	4.59 (0.66, 8.51)
Colombia (2010)	910/6953	13.09 (12.30, 13.88)	NA	NA
Ecuador (2012)	282/4151	6.79 (6.03, 7.56)	NA	NA
UK (GB2014)	33/545	6.06 (4.05, 8.06)	NA	NA
Georgia (2009)	22/178	12.36 (7.52, 17.19)	NA	NA
Laos (2006)	22/170	12.94(7.90. 17.99)	NA	NA
Liberia (2011)	68/378	17.99 (14.12, 21.86)	85/367	23.16 (18.84, 27.48)
Malawi (2016)	112/509	22.00 (18.40, 25.60)	171/503	34.00 (29.86, 38.14)
Mexico (2006)	181/1891	9.54 (8.22, 10.86)	NA	NA
Mexico (2012)	93/1110	8.38 (6.75, 10.01)	NA	NA
PNG (2005)	41/132	31.06 (23.17, 38.95)	NA	NA
US (2006)	332/3246	10.23 (9.19, 11.27)	NA	NA
Vietnam (2010)	9/191	4.71 (1.71, 7.72)	NA	NA
Infection burden				
Low	387	9.75 (8.83, 10.67)	NA	NA
Moderate	1644/15451	10.64 (10.51, 11.13)	NA	NA
High	281/1299	21.63 (19.39, 23.87)	261/979	26.6623.89, 29.43)

1. Inflammation was defined as a CRP concentration >5 mg/L or AGP concentration >1 g/L (only CRP data were available for Colombia, Ecuador, UK, Georgia, Mexico, United States and Vietnam), NA, not available.

Table 3 a: Univariate association between prevalence (%) of anemia by iron and vitamin A deficiencies by country[1]

Country(year)	Iron					Vitamin A				
	Deficient		Sufficient		P-value	Deficient		Sufficient		P-value
	n/N	%	n/N	%		n/N	%	n/N	%	
Azerbaijan (2013)	79/134	58.96	48/227	21.15	<0.0001	2/4	50.00	125/357	35.01	0.61*
Bangladesh (2012)	27/72	37.50	97/711	13.64	<0.0001	38/146	26.03	87/643	13.53	0.0002
Cote d'Ivoire (2007)	15/18	83.33	50/92	54.35	0.02	NA	NA	65/110	59.09	NA
Colombia (2010)	135/1173	11.51	237/5780	4.10	<0.0001	NA	NA	NA	NA	NA
Ecuador (2012)	119/374	31.82	109/3776	2.89	<0.0001	0/10	0.00	0/82	0.00	NA
UK(GB2014)	16/109	14.68	11/416	2.64	<0.0001	0/4	0.00	28/502	5.58	1.00*
Georgia (2009)	NA	NA	45/178	25.28	NA	NA	NA	NA	NA	NA
Laos (2006)	32/61	52.46	37/109	33.94	0.02	NA	NA	NA	NA	NA
Liberia (2011)	86/146	58.90	63/232	27.16	<0.0001	2/5	40.00	147/373	39.41	1.00
Malawi (2016)	21/50	42.00	80/459	17.43	<0.0001	7/23	30.43	94/486	19.34	0.18*
Mexico (2006)	76/476	15.97	113/1408	8.03	<0.001	NA	NA	NA	NA	NA
Mexico (2012)	12/165	7.27	62/941	6.59	0.74	2/12	16.67	54/816	6.62	0.19*
PNG (2005)	NA	NA	NA	NA	NA	NA	NA	54/132	40.91	NA
US (2006)	96/423	22.70	65/1455	4.47	<0.0001	3/11	27.27	178/3141	5.67	0.02
Vietnam (2010)	10/31	32.26	3/160	1.88	<0.0001 *	0/4	0.00	12/180	6.67	1.00*
Infection burden										
Low	112/532	21.05	121/2049	5.91	<0.0001	3/15	20.00	206/3643	5.65	0.05*
Moderate	458/2425	18.89	669/13003	5.14	<0.0001	42/176	23.86	278/2078	13.38	<0.0001
High	154/275	56.00	230/892	25.78	<0.0001	9/28	32.14	360/1101	32.70	0.95

1. Iron deficiency was defined as an inflammation-adjusted ferritin concentration <15 µg/L and vitamin A deficiency was defined as a retinol-binding protein or retinol concentration < 0.7µmol/L; NA, not available.

Table 3 b: Univariate association between prevalence of anemia by folate and vitamin B-12 deficiencies by country[1]

Country(year)	Folate					Vitamin B-12				
	Deficient		Sufficient		P-value	Deficient		Sufficient		P-value
	n/N	%	n/N	%		n/N	%	n/N	%	
Azerbaijan (2013)	55/132	41.67	65/213	30.52	0.03	14/38	36.84	48/135	35.56	0.88
Bangladesh (2012)	18/68	26.47	9/92	9.78	0.01	0/4	0.00	27/160	16.88	1.00*
Cote d'Ivoire (2007)	53/93	56.99	9/13	69.23	0.40	8/11	72.73	25/45	55.56	0.50*
Colombia (2010)	NA	NA	NA	NA	NA	12/137	8.76	126/2828	4.46	0.02
Ecuador (2012)	2/30	6.67	226/4121	5.48	0.69*	6/39	15.38	222/4023	5.52	0.02*
UK(GB2014)	NA	NA	NA	NA	NA	2/20	10.00	26/496	5.24	0.29*
Georgia (2009)	1/20	5.00	2/6	33.33	0.12*	NA	NA	NA	NA	NA
Laos (2006)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Liberia (2011)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Malawi (2016)	11/30	36.67	31/130	23.85	0.15	1/18	5.56	41/142	28.87	0.04
Mexico (2006)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Mexico (2012)	0	0.00	74/1103	6.71	1.00*	1/7	14.29	73/1098	6.65	0.31*
PNG (2005)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
US (2006)	3/26	11.54	182/3200	5.69	0.18*	4/21	19.05	179/3202	5.59	0.03*
Vietnam (2010)	5/23	21.74	6/154	3.90	0.01*	1/3	5.77	2/49	4.08	0.16*
Infection burden										
Low	4/46	8.70	184/3206	5.74	0.34*	6/41	14.63	205/3698	5.54	0.03*
Moderate	80/255	31.37	380/5683	6.69	<0.0001	34/228	14.91	498/8293	6.01	<0.0001
High	64/123	52.03	40/143	27.97	<0.0001	9/29	31.03	66/187	35.29	0.65

1. Folate deficiency was defined as a folate concentration <10 nmol/L. Vitamin B-12 deficiency was defined as a vitamin B-12 concentration <150 pmol/L, Pearson's chi-square P values indicate that the proportion in at least one subgroup is significantly different from the values in the other subgroups, * indicates P-value calculated from Fisher's exact test, NA, not available.

Table 3 c: Univariate association between prevalence of anemia by inflammation and malaria by country

Country(year)	Inflammation[1]					Malaria				
	Any		No		P-value	Yes		No		P-value
	n/N	%	n/N	%		n/N	%	n/N	%	
Azerbaijan (2013)	26/74	35.19	101/287	35.14	0.99	NA	NA	NA	NA	NA
Bangladesh (2012)	19/95	20.00	107/699	15.31	0.24	NA	NA	NA	NA	NA
Cote d'Ivoire (2007)	28/38	73.68	37/72	51.39	0.02	5/5	100.00	59/104	56.73	0.07*
Colombia (2010)	71/910	7.80	301/6043	4.98	<0.01	NA	NA	NA	NA	NA
Ecuador (2012)	27/282	9.57	201/3869	5.20	<0.01	NA	NA	NA	NA	NA
UK(GB2014)	4/33	12.12	24/512	4.69	0.08*	NA	NA	NA	NA	NA
Georgia (2009)	9/22	40.91	36/156	23.08	0.07	NA	NA	NA	NA	NA
Laos (2006)	10/22	45.45	59/148	39.86	0.62	NA	NA	NA	NA	NA
Liberia (2011)	37/68	54.41	112/310	36.13	0.01	37/85	43.53	111/282	39.36	0.49
Malawi (2016)	29/112	25.89	72/397	18.41	0.07	53/171	30.99	47/332	14.16	<0.0001
Mexico (2006)	29/181	16.02	161/1710	9.42	<0.01	NA	NA	NA	NA	NA
Mexico (2012)	10/93	10.75	64/1017	6.29	0.10	NA	NA	NA	NA	NA
PNG (2005)	19/41	46.34	35/91	38.46	0.39	NA	NA	NA	NA	NA
US (2006)	31/332	9.34	154/2914	5.28	<0.01	NA	NA	NA	NA	NA
Vietnam (2010)	1/9	11.11	12/182	6.59	0.47*	NA	NA	NA	NA	NA
Infection burden										
Low	44/387	11.37	214/3582	5.97	<0.0001	NA	NA	NA	NA	NA
Moderate	183/1644	11.13	947/13807	6.86	<0.0001	NA	NA	NA	NA	NA
High	123/281	43.77	315/1018	30.94	<0.0001	95/261	36.40	217/718	30.22	0.07

1. Any inflammation was defined as a CRP concentration >5 mg/L or AGP concentration >1 g/L (only CRP data were available for Colombia, Ecuador, UK, Georgia, Mexico, United States and Vietnam). Pearson's chi-square P values indicate that the proportion in at least one subgroup is significantly different from the values in the other subgroups, * indicates P-value calculated from Fisher's exact test, NA, not available.

Table 3 d: Univariate association between prevalence of anemia by residence and household education attainment by country[1]

Country(year)	Residence					Education Attainment				
	Rural		Urban		P-value	None/Primary		Secondary/University/Trade		P-value
	n/N	%	n/N	%		n/N	%	n/N	%	
Azerbaijan(2013)	81/253	32.02	46/108	42.59	0.05	NA	NA	NA	NA	NA
Bangladesh(2012)	46/263	17.49	80/531	15.07	0.38	83/529	15.69	43/263	16.35	0.81
Cote d'Ivoire(2007)	32/52	61.54	33/58	56.90	0.62	13/15	86.67	50/93	53.76	0.02
Colombia(2010)	136/2382	5.71	236/4571	5.16	0.33	92/1825	5.04	71/1347	5.27	0.77
Ecuador(2012)	91/1720	5.29	137/2431	5.64	0.63	NA	NA	NA	NA	NA
UK(GB2014)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Georgia(2009)	20/67	29.85	25/111	22.52	0.27	NA	NA	NA	NA	NA
Laos(2006)	56/136	41.18	13/34	38.24	0.75	48/117	41.03	20/51	39.22	0.83
Liberia(2011)	72/171	42.11	77/207	37.20	0.33	NA	NA	NA	NA	NA
Malawi(2016)	88/433	20.32	13/76	17.11	0.52	NA	NA	NA	NA	NA
Mexico(2006)	60/664	9.04	130/1227	64.89	0.28	125/1201	10.41	65/680	9.56	0.56
Mexico(2012)	41/679	6.04	33/431	7.66	0.29	14/163	8.59	60/945	6.35	0.29
PNG(2005)	47/109	43.12	7/23	30.43	0.26	NA	NA	NA	NA	NA
US(2006)	NA	NA	NA	NA	NA	12/425	2.82	165/2709	6.09	0.0007
Vietnam(2010)	5/106	4.72	8/85	9.41	0.20	NA	NA	NA	NA	NA
Infection burden										
Low	20/67	29.85	25/111	22.52	0.28	12/425	2.82	165/2709	6.09	0.01
Moderate	460/6067	7.58	670/9384	7.14	0.30	314/3718	8.45	239/3235	7.39	0.10
High	295/901	32.74	143/398	35.93	0.26	48/117	41.03	20/51	39.22	0.82

1. Pearson's chi-square P values indicate that the proportion in at least one subgroup is significantly different from the values in the other subgroups, NA, not available.

Table 3 e: Univariate association between prevalence of anemia by sanitation facility and water source by country[1]

Country(year)	Sanitation Facility					Water Source				
	Unimproved		Improved		P-value	Unimproved		Improved		P-value
	n/N	%	n/N	%		n/N	%	n/N	%	
Azerbaijan (2013)	14/27	51.85	113/334	33.83	0.06	25/84	29.76	102/277	36.82	0.24
Bangladesh (2012)	30/178	16.85	96/616	15.58	0.68	3/18	16.67	123/776	15.85	1.00*
Cote d'Ivoire (2007)	28/49	57.14	37/61	60.66	0.70	13/15	86.67	50/93	53.76	0.02
Colombia (2010)	10/275	3.64	89/1854	4.80	0.39	15/335	4.48	84/1782	4.71	0.85
Ecuador (2012)	14/247	5.67	214/3904	NA	0.90	48/844	5.69	180/3307	5.44	0.78
UK(GB2014)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Georgia (2009)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Laos (2006)	50/97	51.55	19/73	26.03	0.0008	33/82	40.24	36/88	40.91	0.93
Liberia (2011)	79/194	40.72	67/174	38.51	0.66	11/39	28.21	137/338	40.53	0.14
Malawi (2016)	16/83	19.28	85/426	19.95	0.89	18/82	0.60	83/427	19.44	0.60
Mexico (2006)	13/97	13.40	76/567	13.40	0.99	NA	NA	NA	NA	NA
Mexico (2012)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PNG (2005)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
US (2006)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Vietnam (2010)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infection burden[2]										
Low	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Moderate	81/824	9.83	588/7275	8.08	0.08	91/1281	7.10	489/6142	7.96	0.30
High	173/423	40.90	208/734	28.34	<0.0001	75/218	34.40	306/946	32.35	0.56

1. Household sanitation and drinking water source were defined according to the WHO/UNICEF Joint Monitoring Program for water supply and sanitation. Pearson's chi-square Pearson's chi-square P values indicate that the proportion in at least one subgroup is significantly different from the values in the other subgroups, * indicates P-value calculated from Fisher's exact test, NA, not available.
2. Countries were categorized by infection burden as follows—low: Georgia, United Kingdom and the United States; moderate: Colombia and Mexico (2006 and 2012); high: Colombia, Mexico, Ecuador, Vietnam, Azerbaijan and Bangladesh.

Table 3 f: Univariate association between prevalence of anemia by socioeconomic status by country

Country (year)	Socioeconomic Status						
	Low		Mid		High		P-value[2]
	n/N	%	n/N	%	n/N	%	
Azerbaijan (2013)	62/152	40.79	43/142	30.28	22/65	33.85	0.16
Bangladesh (2012)	51/291	17.53	54/334	16.17	21/169	12.43	0.35
Cote d'Ivoire (2007)	27/41	65.85	30/53	56.60	8/16	50.00	0.48
Colombia (2010)	268/4197	6.39	81/2147	3.77	23/609	3.78	<0.0001*
Ecuador (2012)	113/1991	5.68	88/1558	5.65	27/601	4.49	0.51
UK(GB2014)	12/207	5.80	7/154	4.55	4/133	3.01	0.49
Georgia (2009)	NA	NA	NA	NA	NA	NA	NA
Laos (2006)	35/69	50.72	26/72	36.11	8/29	27.59	0.06
Liberia (2011)	35/85	41.18	76/168	45.24	38/125	30.40	0.03*
Malawi (2016)	36/177	20.34	46/226	20.35	19/106	17.92	0.86
Mexico (2006)	98/982	9.98	68/695	9.78	24/209	11.48	0.77
Mexico (2012)	41/502	8.17	26/462	5.63	7/146	4.79	0.18
PNG (2005)	22/48	45.83	22/51	43.14	10/33	30.30	0.35
US (2006)	109/1437	7.59	52/1193	4.36	10/33	30.30	0.35
Vietnam (2010)	NA	NA	NA	NA	NA	NA	NA
Infection burden[1]							
Low	121/1644	7.36	59/1347	4.38	22/624	3.53	<0.0001
Moderate	633/8115	7.80	360/5338	6.74	124/1799	6.89	0.05
High	155/420	36.90	200/570	35.09	83/309	26.86	0.01

1. Countries were categorized by infection burden as follows— low: Georgia, United Kingdom and the United States; moderate: Colombia, Mexico (2006 and 2012), Ecuador, Vietnam, Azerbaijan, Bangladesh; high: Côte d'Ivoire, Liberia, Laos, and Papua New Guinea, Malawi.
2. Pearson's chi-square P values indicate that the proportion in at least one subgroup is significantly different from the values in the other subgroups, * indicates P-value calculated from Fisher's exact test, NA, not available.

Table 4 Univariate logistic regression for pooled analysis

			Anemia		
Covariate	**Level**	**N**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Vitamin A	Deficient	219	2.32 (1.69-3.18)	**<.001**	**<.001**
	Sufficient	6822	-	-	
Iron	Deficient	3232	4.22 (3.81-4.69)	**<.001**	**<.001**
	Sufficient	15944	-	-	
Plasma/Serum folate	Deficient	424	7.48 (6.03-9.29)	**<.001**	**<.001**
	Sufficient	9032	-	-	
Serum Vitamin B12	Deficient	298	2.92 (2.13-4.00)	**<.001**	**<.001**
	Sufficient	12178	-	-	
Inflammation	Any	2312	2.05 (1.80-2.32)	**<.001**	**<.001**
	No	18407	-	-	
Malaria	Yes	261	1.32 (0.98-1.78)	0.067	0.067
	No	718	-	-	
Sex	Female	15267	2.38 (2.08-2.73)	**<.001**	**<.001**
	Male	5452	-	-	
Socioeconomic Status	Low	10179	1.07 (0.92-1.25)	0.370	0.528
	Medium	7255	1.02 (0.87-1.19)	0.811	
	High	2732	-	-	
Residence Type	rural	7035	1.34 (1.21-1.48)	**<.001**	**<.001**
	urban	9893	-	-	
Toilet Source	Unimproved	1247	2.32 (1.98-2.71)	**<.001**	**<.001**
	Improved	8009	-	-	
Water Source	Unimproved	1499	0.99 (0.83-1.18)	0.874	0.874
	Improved	7088	-	-	
Infection Burden	Moderate	15451	1.13 (0.99-1.31)	0.076	**<.001**
	High	1299	7.32 (6.17-8.68)	**<.001**	
	low	3969	-	-	
Education	Low	4260	1.26 (1.09-1.46)	**0.002**	**0.002**
	High	5995	-	-	
Age in Years		20719	1.21 (1.19-1.23)	**<.001**	**<.001**

Table 5 Univariate logistic regression in low infection burden

			Anemia		
Covariate	**Level**	**N**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Vitamin A	Deficient	15	4.17 (1.17-14.90)	**0.028**	**0.028**
	Sufficient	3643	-	-	
Iron	Deficient	532	4.25 (3.22-5.61)	**<.001**	**<.001**
	Sufficient	2049	-	-	
Plasma/Serum folate	Deficient	46	1.56 (0.55-4.41)	0.398	0.398
	Sufficient	3206	-	-	
Serum Vitamin B12	Deficient	41	2.92 (1.21-7.02)	**0.017**	**0.017**
	Sufficient	3698	-	-	
Inflammation	Any	387	2.02 (1.43-2.84)	**<.001**	**<.001**
	No	3582	-	-	
Sex	Female	2756	5.07 (3.26-7.89)	**<.001**	**<.001**
	Male	1213	-	-	
SES	Low	1644	2.17 (1.37-3.46)	**0.001**	**<.001**
	Medium	1347	1.25 (0.76-2.06)	0.375	
	High	624	-	-	
Residence Type	Rural	67	1.46 (0.74-2.91)	0.277	0.277
	Urban	111	-	-	
Education	Low	425	0.45 (0.25-0.81)	**0.008**	**0.008**
	High	2709	-	-	
Age in Years		3969	1.27 (1.21-1.33)	**<.001**	**<.001**
Malaria		0	NA	NA	NA
Water Source		0	NA	NA	NA
Toilet Source		0	NA	NA	NA

Table 6 Univariate logistic regression in moderate infection burden

			Anemia		
Covariate	**Level**	**N**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Vitamin A	Deficient	176	2.03 (1.40-2.93)	**<.001**	**<.001**
	Sufficient	2078	-	-	
Iron	Deficient	2425	4.29 (3.78-4.88)	**<.001**	**<.001**
	Sufficient	13003	-	-	
Plasma/Serum folate	Deficient	255	6.38 (4.80-8.48)	**<.001**	**<.001**
	Sufficient	5683	-	-	
Serum Vitamin B12	Deficient	228	2.74 (1.88-3.99)	**<.001**	**<.001**
	Sufficient	8293	-	-	
Inflammation	Any	1644	1.70 (1.44-2.01)	**<.001**	**<.001**
	No	13807	-	-	
Sex	Female	11394	1.78 (1.52-2.09)	**<.001**	**<.001**
	Male	4057	-	-	
SES	Low	8115	1.14 (0.94-1.40)	0.190	0.054
	Medium	5338	0.98 (0.79-1.21)	0.828	
	High	1799	-	-	
Residence Type	Rural	6067	1.07 (0.94-1.21)	0.303	0.303
	Urban	9384	-	-	
Toilet Source	Unimproved	824	1.24 (0.97-1.58)	0.085	0.085
	Improved	7275	-	-	
Water Source:	Unimproved	1281	0.88 (0.70-1.12)	0.298	0.298
	Improved	6142	-	-	
Education	Low	3718	1.16 (0.97-1.38)	0.104	0.104
	High	3235	-	-	
Age in Years		15451	1.15 (1.13-1.18)	**<.001**	**<.001**
Malaria		0	NA	NA	NA

Table 7 Univariate logistic regression in high infection burden

			Anemia		
Covariate	**Level**	**N**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Vitamin A	Deficient	28	0.98 (0.44-2.18)	0.951	0.951
	Sufficient	1101	-	-	
Iron	Deficient	275	3.66 (2.76-4.85)	**<.001**	**<.001**
	Sufficient	892	-	-	
Plasma/Serum folate	Deficient	123	2.79 (1.68-4.64)	**<.001**	**<.001**
	Sufficient	143	-	-	
Serum Vitamin B12	Deficient	29	0.83 (0.36-1.91)	0.654	0.654
	Sufficient	187	-	-	
Inflammation	Any	281	1.74 (1.33-2.28)	**<.001**	**<.001**
	No	1018	-	-	
Sex	Female	1117	2.37 (1.61-3.50)	**<.001**	**<.001**
	Male	182	-	-	
SES	Low	420	1.59 (1.16-2.19)	**0.004**	**0.012**
	Medium	570	1.47 (1.09-2.00)	**0.013**	
	High	309	-	-	
Residence Type	Rural	901	0.87 (0.68-1.11)	0.263	0.263
	Urban	398	-	-	
Toilet Source	Unimproved	423	1.75 (1.36-2.25)	**<.001**	**<.001**
	Improved	734	-	-	
Water Source	Unimproved	218	1.10 (0.80-1.50)	0.560	0.560
	Improved	946	-	-	
Education	Low	117	1.08 (0.55-2.11)	0.827	0.827
	High	51	-	-	
Malaria	Yes	261	1.32 (0.98-1.78)	0.067	0.067
	No	718	-	-	
Age in Years		1299	1.19 (1.14-1.25)	**<.001**	**<.001**

Table 8 Multivariable logistic model in low infection burden group

		Anemia		
Covariate	**Level**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Vitamin A	Deficient	13.84 (1.58-120.91)	**0.017**	**0.017**
	Sufficient	-	-	
Iron	Deficient	6.08 (4.26-8.68)	**<.001**	**<.001**
	Sufficient	-	-	
Inflammation	Any	1.88 (1.16-3.03)	**0.010**	**0.010**
	No	-	-	
SES	Low	2.07 (1.17-3.68)	**0.013**	**0.008**
	Medium	1.27 (0.69-2.34)	0.444	
	High	-	-	
Education	Low	0.30 (0.14-0.61)	**<.001**	**<.001**
	High	-	-	
* Number of observations in the original data set = 3969. Number of observations used = 1708.				

Table 9 Multivariable logistic model in moderate infection burden group

		Anemia		
Covariate	**Level**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Plasma/Serum folate	Deficient	2.59 (1.87-3.59)	**<.001**	**<.001**
	Sufficient	-	-	
Iron	Deficient	7.20 (5.81-8.93)	**<.001**	**<.001**
	Sufficient	-	-	
Inflammation	Deficient	1.91 (1.40-2.60)	**<.001**	**<.001**
	Sufficient	-	-	
Age in Years		1.20 (1.15-1.25)	**<.001**	**<.001**
* Number of observations in the original data set = 15451. Number of observations used = 5936.				

Table 10 Multivariable logistic model in high infection burden group

		Anemia		
Covariate	**Level**	**Odds Ratio (95% CI)**	**OR P-value**	**Type3 P-value**
Plasma/Serum folate (SFO)	Deficient	2.48 (1.46-4.18)	**<.001**	**<.001**
	Sufficient	-	-	
Iron	Deficient	2.45 (1.29-4.64)	**0.006**	**0.006**
	Sufficient	-	-	
Inflammation	Any	2.14 (1.18-3.87)	**0.012**	**0.012**
	No	-	-	
* Number of observations in the original data set = 1299. Number of observations used = 266.				

Supplement Table 1: Criteria of categorize as infection burden group

	Malaria intensity[1]		HIV prevalence in adults[3]		Drinking water quality[4]		Sanitary situation[4]		Overall hygiene score	Schistosomiasis prevalence[5]		Total points	Infection categories
Country	**Presumed and confirmed malaria cases per 100 in 2013**		**Proportion of adults with HIV/Aids 2017**		**Proportion of population using improved drinking water sources (%), 2015**		**Proportion of population using improved sanitation facilities (%) 2015**		**Average of drinking water quality**	**Distribution of Schistosomiasis (%)**			
	(0-<5% = 0; 5%-<15%= 1; 15%-<25%= 2; ≥ 25%=3)		**(0= <1%; 1= 1-9.9%; 2= 10-19.9%; 3= ≥ 20%)**		**(>90%= 0; 76-90%= 1; 75-50%= 2; < 50%= 3)**		**(>90%= 0; 76-90%= 1; 75-50%= 2; < 50%= 3)**		**and sanitation situation**	**(0%= 0; <10%= 1; 10-49%= 2; > 50%=3)**		**Maximum 12 points**	**Low = 0-0.49; Medium= 0.5-2.9; High= ≥3**
	Proportion (%)	**Category**	**Proportion (%)**	**Category**	**Proportion (%)**	**Category**	**Proportion (%)**	**Category**	**Score**	**Proportion (%)**	**Category**		**Category**
Cameroon	23	2	3.7	1	50-75	2	<50	3	2.5	10-49	2	7.5	High
Colombia	<1	0	0.5	0	> 90	0	76-90	1	0.5	0	0	0.5	Medium
Cote d'Ivoire	37	3	2.8	1	76-90	1	<50	3	2.0	10-49	2	8.0	High
Georgia	<1	0	0.4	0	> 90	0	>90	0	0.0	0	0	0.0	Low
Laos	1	0	0.3	0	50-75	2	50-75	2	2.0	<10	1	3.0	High
Liberia	36	3	1.4	1	50-75	2	<50	3	2.5	10-49	2	8.5	High
Mexico	<1	0	0.3	0	> 90	0	76-90	1	0.5	0	0	0.5	Medium
Papua New Guinea	17	2	0.9	0	<50	3	<50	3	3.0	0	0	5.0	High
United States	0*	0	<0.1	0	>90	0	>90	0	0.0	0	0	0.0	Low
UK	0*	0	<0.1	0	>90	0	>90	0	0.0	0	0	0.0	Low
Ecuador	<1	0	0.3	0	76-90	1	76-90	1	1.0	0	0	1.0	Medium
Afghanistan	<1	0	<0.1	0	76-90	1	<50	3	2.0	0	0	2.0	Medium
Vietnam	2	0	0.3	0	76-90	1	>90	1	1.0	0	0	1.0	Medium
Malawi	14	1	9.6	1	76-90	1	<50	3	2.0	10-49	2	4.0	High
Azerbaijan	<1	0	0.1	0	76-90	1	76-90	1	1.0	0	0	1.0	Medium
Bangladesh	<1	0	0.1	0	76-90	1	50-75	2	1.5	0	0	1.5	Medium
Pakistan	4	0	0.1	0	>90	0	50-75	2	1.0	0	0	1.0	Medium

*Non malaria endemic country[2]

Sources:	
[1] WHO World Malaria Report 2015	http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/
[2] Malaria Atlas Project	http://www.map.ox.ac.uk/explore/countries/
3 World Bank, prevalence of HIV	https://data.worldbank.org/indicator/SH.DYN.AIDS.ZS?name_desc=false
[4] WHO 2015, Proportion of population using improved drinking water sources/sanitation facilities (2015)	http://gamapserver.who.int/mapLibrary/
[5] WHO 2012; Prevalence of Schistosomiasis	http://gamapserver.who.int/mapLibrary/Files/Maps/Schistosomiasis_2011_global.png

Supplement table 2:
Likelihood Ratio Statistics for Type 3 Analysis with 12 Main Effect and Interaction Models (Assess interaction between covariates and infection burden group)

Covariate	DF	Chi-Square	P-value	Sig	Covariate	DF	Chi-Square	P-value	Sig
Infection group	2	15.23	0.0005	**	Infection group	2	181.62	<0.0001	***
vitamin A deficiency	1	5.58	0.02	*	sex	1	126.53	<0.0001	***
Interaction	2	4.16	0.13		Interaction	2	23.48	**<0.0001**	***
Infection group	2	469.25	<0.0001	***	Infection group	2	586.98	<0.0001	***
Iron deficiency	1	372.87	<0.0001	***	SES status	2	26.43	<0.0001	***
Interaction	2	1.02	0.60		Interaction	4	13.20	**0.01**	*
Infection group	2	94.82	<0.0001	***	Infection group	2	609.46	<0.0001	***
Folate deficiency	1	17.21	<0.0001	***	Rural/urban	1	0.65	0.42	
Interaction	2	14.00	**0.0009**	**	interaction	2	3.13	0.21	
Infection group	2	31.87	<0.0001	***	Infection group	1	362.73	<0.0001	***
Vitamin B12 deficiency	1	7.29	0.0069	**	Sanitation	1	17.97	<0.0001	***
Interaction	2	7.19	**0.03**	*	interaction	1	3.77	**0.05**	*
Infection group	2	456.62	<0.0001	***	Infection group	1	285.15	<0.0001	***
Inflammation	1	49.74	<0.0001	***	Water	1	0.02	0.88	
Interaction	2	0.77	0.68		interaction	1	1.19	0.28	
Infection group	2	181.62	<0.0001	***	Infection group	2	129.93	<0.0001	***
sex	1	126.53	<0.0001	***	Education	1	1.59	0.21	
Interaction	2	23.48	<0.0001	***	Interaction	2	11.09	**0.0039**	**

***p<0.0001, **p<0.01, *p<0.05