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:

_____ _____

Monica Shah Date

Effect of transmission reduction by ITNs on the prevalence of mutations associated with resistance to sulfadoxine-pyrimethamine and chloroquine in western Kenya

By

Monica Shah
Master of Public Health

Department of Global Epidemiology

_____

Kevin Sullivan, PhD, MPH, MHA
Committee Chair (Faculty Thesis Advisor)

_____

Ya Ping Shi, MD, MSc
Committee Member (Thesis Field Advisor)

Effect of transmission reduction by ITNs on the prevalence of mutations associated with resistance to sulfadoxine-pyrimethamine and chloroquine in western Kenya

By

Monica Shah
B.S. University of North Carolina at Chapel Hill, 2008

Faculty Thesis Advisor: Kevin Sullivan, PhD, MPH, MHA
Thesis Field Advisor: Ya Ping Shi, MD, MSc

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 Public Health
in the Department of Global Epidemiology
2011

Abstract

Effect of transmission reduction by ITNs on the prevalence of mutations associated with resistance to sulfadoxine-pyrimethamine and chloroquine in western Kenya

By Monica Shah

Background: Malaria is a devastating disease affecting people living in tropical areas, particularly in sub-Saharan Africa. Despite the clear benefit of insecticide-treated bednets (ITNs) in preventing malaria infection, the impact of malaria transmission-reduction by vector control on the spread of drug resistance is not well understood. We investigated the effect of sustained transmission reduction by ITNs on the prevalence of *Plasmodium falciparum* drug resistant gene mutations in an ITN trial carried out between 1996 and 2001 in western Kenya during a national drug policy shift from chloroquine (CQ) to sulfadoxine-pyrimethamine (SP).

Methods: We compared the prevalence of mutations at *dhfr*-51,59,108,164 and *dhps*-437, 540 (SP resistance) and *pfcrt*-76 and *pfmdr1*-86 (CQ resistance) in *P. falciparum* smear-positive samples collected from children under the age of five years during cross-sectional surveys prior to ITN introduction (baseline, n=250) and five years post-ITN intervention (year 5 survey, n=242). Multivariable logistic regression models were used to explore the association between two primary exposures of interest, survey year and antimalarial drug use (antifolate class, SP, or CQ), and drug resistance genotypes.

Results: We observed significant increases in the prevalence of *dhps* mutations and the SP quintuple mutant ($p<0.0001$), and a significant reduction in the proportion of mixed infections detected at *dhfr*-51,59 and *dhps*-437,540 SNPs ($p<0.004$) from baseline to the year 5 survey. There was no change in the high prevalence of CQ mutations (82% and 75% at baseline to 82% and 73% at year 5 survey, for *pfcrt-76* and *pfmdr1-86*, respectively). Multivariable regression results showed that antifolate drug use (*dhps* mutations aOR, 2.4 [95% CI, 1.2-5.1]) and year of survey (*dhps* mutations aOR, 10.3 [95% CI, 6.2-17.2]; *dhfr/dhps* mutations aOR, 8.8 [95% CI, 5.5-14.3]) were significantly associated with more SP drug resistant mutations.

Conclusions: Our results suggest that increased antifolate use likely led to the high prevalence of SP drug resistant mutations 5 years post-ITN intervention and reduced transmission had no apparent effect on the existing high prevalence of CQ drug resistant mutations. There is no evidence from the current study that sustained transmission reduction by ITNs reduces the prevalence of genes associated with antimalarial drug resistance.

Effect of transmission reduction by ITNs on the prevalence of mutations associated with resistance to sulfadoxine-pyrimethamine and chloroquine in western Kenya

By

Monica Shah
B.S. University of North Carolina at Chapel Hill, 2008

Faculty Thesis Advisor: Kevin Sullivan, PhD, MPH, MHA
Thesis Field Advisor: Ya Ping Shi, MD, MSc

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 Public Health

in the Department of Global Epidemiology

2011

ACKNOWLEDGEMENTS

I am most indebted to Dr. Ya Ping Shi, my thesis field advisor, and Dr. Kevin Sullivan, my faculty thesis advisor, for their guidance and advice. This research would have not been possible without the contributions and feedback of my co-authors, Dr. Simon Kariuki, Jodi Vanden Eng, Anna J Blackstock, Kimberly Garner, Dr. Wangeci Gatei, Dr. John E Gimnig, Dr. Kim Lindblade, Dr. Dianne Terlouw, Dr. Feiko ter Kuile, Dr. William A Hawley, Dr. Penelope Phillips-Howard, Dr. Bernard Nahlen, Dr. Edward Walker, Dr. Mary J Hamel, Dr. Laurence Slutsker and Dr. Ya Ping Shi. I also wish to extend gratitude to the children and their caretakers who participated in the ITN trial. I am grateful to the CDC/KEMRI field workers and malaria laboratory staff in Kisumu and Atlanta who assisted with the ITN trial and this study. Finally, a special thanks to my family for their encouragement, love, and support along my academic path.

TABLE OF CONTENTS

ACRONYMS

ACT	Artemisinin-based combination therapy
AL	Artemether-Lumefantrine
aOR	Adjusted odds ratio
BX	Bednet cross-sectional survey
CDC	Centers for Disease Control and Prevention
CI	Confidence interval
CQ	Chloroquine
CTX	Co-trimoxazole, trimethoprim/sulfamethoxazole (septrin)
DHFR	Dihydrofolate reductase
DHPS	Dihydropteroate synthase
EIR	Entomologic inoculation rate
GIS	Geographic information system
HIV	Human immunodeficiency virus
ITC	Insecticide-treated curtain
ITN	Insecticide-treated bednet
KEMRI	Kenya Medical Research Institute
LLITN	Long lasting insecticide-treated bednets
MDG	Millennium Development Goal
MS	Microsatellite
OR	Odds ratio
PCR	Polymerase chain reaction
PFCRT	Plasmodium falciparum chloroquine-related transporter
PFMDR1	Plasmodium falciparum multidrug resistance gene
SD	Standard deviation
SNP	Single nucleotide polymorphism
SP	Sulfadoxine-pyrimethamine
VDP	Variance decomposition proportion
WHO	World Health Organization

CHAPTER I: BACKGROUND/LITERATURE REVIEW

Burden of Malaria

Globally, 3 billion people are at risk of malaria infection, a parasitic disease transmitted by the bite of an infective *Anopheles* mosquito. Although the parasite thrives in tropical environments worldwide, the greatest burden of disease occurs in Africa. In 2009, approximately 79% of the 225 million malaria cases and 91% of the 1 million malaria deaths occurred in Africa. Children under five years of age bear the brunt of the disease burden, as 85% of all malaria deaths occur in this age group. Of the five *Plasmodium* parasite species known to cause malaria in humans, *P. falciparum* is considered the most severe, accounting for 98% of deaths due to malaria in sub-Saharan Africa alone (1).

Uncomplicated malaria is characterized by symptoms such as fever, chills, sweats, headaches, nausea and vomiting, while complicated (severe) malaria manifests as severe anemia, cerebral malaria, acute kidney failure, and other critical conditions caused by organ failure. With prompt diagnosis, both conditions can be effectively treated (2). In addition to the adverse individual health consequences of malaria, the disease exerts a tremendous social and economic burden on countries due to factors such as loss of work force time and school absenteeism (3).

In Kenya, approximately 74% of the population is at risk of malaria infection and *P. falciparum* is the predominant species. In 2006, it was estimated that 20% of deaths in children under the age of five were due to malaria (4).

Malaria Control

In the absence of an effective vaccine, malaria is a disease that is primarily managed by prevention tools and treatment. The change in first line drug policy for malaria treatment has mainly been driven by the emergence and spread of drug resistance. *P. falciparum* resistance to chloroquine (CQ) in sub-Saharan Africa was first documented in Kenya in the 1979 (5), however, a national drug policy shift to a new drug, sulfadoxine-pyrimethamine (SP), did not occur until 1998 (6). Recently, due to decreasing the efficacy of SP globally, many countries have adopted artemisinin-based combination therapy (ACT) as first line treatment for uncomplicated malaria. In Kenya, artemether-lumefantrine (AL) officially replaced SP as first-line antimalarial drug policy in 2004 and the policy change became implemented in 2006 (6). As HIV is prevalent in many African countries, prophylaxis using co-trimoxazole (CTX) has been recommended to prevent opportunistic infections in HIV-positive individuals. Similar to SP, CTX is an antifolate drug that provides some protection against malaria (7).

In recent years, the international community has placed a greater emphasis on malaria prevention in order to preserve the efficacy of the limited set of anti-malarial treatment drugs and subsequently limit the emergence and spread of drug resistant parasites (7, 8). The World Health Organization (WHO) recommends four malaria control strategies to reduce the global burden of disease: prompt access to effective treatment, vector control with long lasting insecticide-treated bednets (LLITNs) and indoor residual spraying (IRS), and prevention of malaria in pregnancy (1). Insecticide-treated bednets (ITNs) are considered a cost-effective intervention and the use of ITNs has become widespread in highly endemic areas (9). Studies in sub-Saharan Africa have shown that ITN use was

associated with a 70-90% decrease in malaria transmission, reduction in childhood malaria morbidity and all-cause mortality, and significant decrease in adverse effects of malaria in pregnancy (9-11).

Transmission intensity and the spread of drug resistance: conceptual frameworks

By repelling and killing infective *anopheles* mosquitoes, ITNs decrease the transmission of malaria and, subsequently, reduce infections. Although the public health benefit of ITNs has been well-described, the effect of transmission reduction by ITN use on the spread of antimalarial drug resistance is less clear. It is believed that by decreasing transmission, ITNs may ultimately reduce the number of drug exposed parasites; however, the relationship between transmission intensity and the spread of drug resistance is complex and relies on parasite, human host, and genetic factors.

Recent conceptual frameworks describe the role of transmission intensity on the spread of drug resistance as indirect, where the intensity of transmission affects three main epidemiological factors ("mediators") that modulate five effector variables, and these "effectors" directly shape the dynamics of drug resistance (Figure 1). The average number of malaria infections per human host (clonal multiplicity) modulates the effectors sexual recombination and intrahost dynamics, the threat of infection determines the level of drug use in the population, and host immunity affects the proportion of malaria infections treated and the number of parasites in a human host ("biomass") (8, 12).

When considered separately, clonal multiplicity, more specifically defined as "the number of independently acquired and infective parasite clones circulating within a single host at a given time," (13) plays an important role in the spread of drug resistance.

As *Plasmodium* parasites undergo an obligate sexual phase, the genetic material carried by gametocytes will determine whether drug resistant genes are spread within a population. If gametocytes from multiple clones are taken up by the mosquito during a blood meal, genetic reassortment can occur, leading to either the creation or break down of the gene combinations encoding drug resistance to one or more drugs (13). However, in areas of high transmission where clonal multiplicity is greater, the effectors sexual recombination and intrahost dynamics could have antagonistic effects on the evolution of drug resistance. While sexual recombination may break down the combination of gene mutations required to confer drug resistance, the presence of intrahost dynamics may facilitate the spread of resistant clones after clearance of sensitive clones by treatment.

Transmission reduction could decrease the threat of malaria infection and change the level of host immunity to disease, the second and third mediators in the model described above, by decreasing the prevalence of malaria infections. As a result, care seeking and treatment behaviors could modulate drug pressure, leading to increased or decreased spread of drug resistance. Fewer malaria infections may lead to fewer febrile illnesses and, subsequently, less presumptive treatment with antimalarial drugs within a community. Reduced community drug use would decrease residual drug levels and limit the spread of drug resistance. On the other hand, as transmission intensity decreases, therapeutic drug use may increase, since reduced immunity may result in an increase in severe and symptomatic malaria infections and a greater proportion of infections being treated with antimalarial drugs (8).

Therapeutic modes of action of antimalarial drugs and mechanisms of drug resistance

CQ is considered a quinolone-containing antimalarial drug. It is believed that CQ interferes with the heme degradation process in red blood cells and therefore disrupts the parasite life cycle during hemoglobin digestion. Mutations in multiple genes play a role in CQ resistance, although the exact mechanisms leading to resistance remains unclear. It is speculated that mutations interfere with drug transport and reduced concentration in the parasite digestive vacuole (14).

SP belongs to a class of folate antagonist drugs that inhibit enzymes in the *P. falciparum* folate pathway, ultimately resulting in reduced parasite DNA synthesis. Drug resistance to SP develops as a result of the accumulation of step-wise mutations in genes encoding enzymes involved in the parasite folic acid pathway. It is hypothesized that these mutations alter the structure of the enzyme's active site such that the drug can no longer bind and inhibit enzyme action (14). In this study, we also consider the combination drug trimethoprim/sulfamethoxazole or co-trimoxazole (CTX), which is widely prescribed in Africa to prevent and treat bacterial infections. CTX inhibits the same enzymes in the folic acid biosynthetic pathway as pyrimethamine and, therefore, may select for similar mutations (15).

Molecular Markers for Antimalarial Drug Resistance

Drug resistance develops through a series of sequential mutations in parasite genes that synergistically confer resistance. The genetic basis of drug resistance to SP is believed to be monogenic, or caused by mutations in a single gene. Point mutations in *P. falciparum* dihydrofolate reductase (*dhfr)* are associated with resistance to

pyrimethamine, while mutations in dihydropteroate synthase (*dhps*) are associated with sulfadoxine resistance. In Africa, the *dhfr* triple mutant, Asn-108/Ile-51/Arg-59, has been strongly associated with clinical resistance to SP and the addition of the *dhps* double mutant, Gly-437/Glu-540, creates the quintuple mutant that is associated with *in vivo* SP treatment failure (16). In contrast, the genetic basis of resistance to CQ is considered multigenic or conferred by mutations in multiple genes. Mutations at Tyr-86 in multidrug resistance gene (*pfmdr1)* and at Thr-76 in chloroquine-related transporter (*pfcrt)* are associated with CQ resistance (17).

Previous Studies

Few studies have sought to assess the effect of vector control interventions on the spread of malaria drug resistance and the studies to date have provided conflicting results. In Tanzania, a study demonstrated that short-term use of ITNs was associated with decreased prevalence of the *dhfr* triple mutant (18). After an IRS campaign conducted in Zimbabwe, participants in sprayed villages had a lower risk of CQ treatment failure and lower prevalence of gene mutations in the parasites conferring resistance to CQ compared to unsprayed villages (19). However, a study examining the effect of long-term use of insecticide-treated curtains (ITCs) on antimalarial drug resistance in Burkina Faso revealed no changes in the risk of CQ treatment failure and prevalence of gene mutations linked to CQ and SP in intervention compared to control villages (20). The results of these studies suggest that decreased malaria transmission by vector control may reduce the number of people exposed to drug resistant parasites and the number of people seeking antimalarial treatment, thereby reducing drug pressure.

CHAPTER II. MANUSCRIPT

Effect of transmission reduction by ITNs on the prevalence of mutations associated with resistance to sulfadoxine-pyrimethamine and chloroquine in western Kenya

Monica Shah

Abstract

Background: Malaria is a devastating disease affecting people living in tropical areas, particularly in sub-Saharan Africa. Despite the clear benefit of insecticide-treated bednets (ITNs) in preventing malaria infection, the impact of malaria transmission-reduction by vector control on the spread of drug resistance is not well understood. We investigated the effect of sustained transmission reduction by ITNs on the prevalence of *Plasmodium falciparum* drug resistant gene mutations in an ITN trial carried out between 1996 and 2001 in western Kenya during a national drug policy shift from chloroquine (CQ) to sulfadoxine-pyrimethamine (SP).

Methods: We compared the prevalence of mutations at *dhfr*-51,59,108,164 and *dhps*-437, 540 (SP resistance) and *pfcrt*-76 and *pfmdr1*-86 (CQ resistance) in *P. falciparum* smear-positive samples collected from children under the age of five years during cross-sectional surveys prior to ITN introduction (baseline, n=250) and five years post-ITN intervention (year 5 survey, n=242). Multivariable logistic regression models were used to explore the association between two primary exposures of interest, survey year and antimalarial drug use (antifolate class, SP, or CQ), and drug resistance genotypes.

Results: We observed significant increases in the prevalence of *dhps* mutations and the SP quintuple mutant ($p<0.0001$), and a significant reduction in the proportion of mixed infections detected at *dhfr*-51,59 and *dhps*-437,540 SNPs ($p<0.004$) from baseline to the year 5 survey. There was no change in the high prevalence of CQ mutations (82% and 75% at baseline to 82% and 73% at year 5 survey, for *pfcrt-76* and *pfmdr1-86*, respectively). Multivariable regression results showed that antifolate drug use (*dhps* mutations aOR, 2.4 [95% CI, 1.2-5.1]) and year of survey (*dhps* mutations aOR, 10.3 [95% CI, 6.2-17.2]; *dhfr/dhps* mutations aOR, 8.8 [95% CI, 5.5-14.3]) were significantly associated with more SP drug resistant mutations.

Conclusions: Our results suggest that increased antifolate use likely led to the high prevalence of SP drug resistant mutations 5 years post-ITN intervention and reduced transmission had no apparent effect on the existing high prevalence of CQ drug resistant mutations. There is no evidence from the current study that sustained transmission reduction by ITNs reduces the prevalence of genes associated with antimalarial drug resistance.

Introduction

Malaria is a devastating disease affecting people living in tropical areas, particularly in sub-Saharan Africa. Annually, almost 1 million people die from malaria which can be prevented using mosquito control tools such as ITNs, and treated using antimalarial drugs (1). Despite the benefit of ITNs in preventing malaria, it is unclear whether reduction of malaria transmission affects the spread of antimalarial drug resistance, as the use of ITNs

changes the parasite-mosquito relationship and human response to malaria infection, ultimately influencing drug pressure (8).

A better understanding of the effect of transmission intensity on the spread of genes that confer resistance to antimalarial drugs could have important public health implications, as the impact of transmission-reducing interventions may award additional benefits if reduced transmission decreases the spread of drug resistance. In the present study, we investigated the impact of sustained malaria transmission reduction by ITNs during first line national drug policy change from CQ to SP on the prevalence of *Plasmodium falciparum* gene mutations associated with drug resistance to SP and CQ in children under the age of five during a large bednet trial conducted between 1996 and 2001 in western Kenya. Using genetic, clinical, and epidemiological data, we 1) determined the prevalence of gene mutations associated with SP and CQ drug resistance before and 5 years after ITN intervention and 2) examined the association between epidemiological variables and mutations in the SP and CQ-linked drug resistance genes.

Methods

Study Site and Population

This study was part of a two-phase ITN trial carried out by the Kenya Medical Research Institute (KEMRI) and US Centers for Disease Control and Prevention (CDC) between 1996 and 2001 in Asembo, western Kenya, where malaria transmission is holoendemic. Detailed methods for the ITN trial are explained elsewhere (21, 22), but briefly described here (Figure 2). During the ITN trial, biannual population censuses and

annual cross-sectional surveys were conducted in 60 villages between March and May (rainy season) to determine the impact of ITNs on malaria-related morbidity and all-cause mortality in children under the age of five. At each cross-sectional survey, blood samples were collected, and parasitological, clinical, demographic, and entomological information were recorded. *P. falciparum* accounted for approximately 98% of malaria infections in the trial area (21). The entomologic inoculation rate (EIR), a measure of transmission intensity, was recorded at 61.3 and 1.3 infective bites per person per year at baseline and five years after ITN introduction, respectively (21, 23). Malaria parasite prevalence in children under the age of five decreased from 70% prior to the ITN trial to 34% at the year 5 survey (21, 24). ITN usage in children younger than five years of age increased from <5% at baseline to 82.5% at the year 5 survey (21). Overall, the number of people seeking antimalarial treatment decreased after the introduction of ITNs (24, 25). During the ITN trial, Kenya's national first line treatment for uncomplicated malaria in children shifted from CQ to SP in 1998. Prior to the policy change and implementation of ITNs (year 1996), SP was sporadically available in health facilities and prescribed occasionally (<1%) in the study area (25) . In addition, we considered CTX as a drug use variable as CTX inhibits the same enzymes in the folic acid biosynthetic pathway as pyrimethamine and, therefore, may select for mutations in *dhfr*. It is speculated that cross-resistance between CTX and SP use, particularly a concern in HIV-infected individuals, may accelerate the spread of resistance to SP by selecting for mutations in *dhfr* and *dhps* (15). The change from Penicillin to CTX for first line treatment of respiratory illnesses occurred in the mid-late 1990s in Kenya.

For this study, 259 *P. falciparum* smear-positive blood samples collected from children under the age of five just prior to ITN introduction (year 1996) and 244 samples collected five years post-intervention as year 5 survey (year 2001) were randomly selected in the same subset of villages. These samples were genotyped for single nucleotide polymorphisms (SNPs) in *P. falciparum* genes related to antimalarial drug resistance. Genotyping was unsuccessful for 9 samples at baseline and 2 samples at the year 5 survey. In this analysis, the molecular marker results at *dhfr-51,59,108,164, dhps-437,540, pfcrt-76,* and *pfmdr1-86* for 250 (96.5%) samples at baseline and 242 (99.2%) samples at the year five survey were analyzed along with clinical, epidemiological, and genetic characteristics (Figure 2).

The original study was approved by the Ethical Review Committee of the KEMRI, Nairobi, Kenya and the Institutional Review Board of the Centers for Disease Control and Prevention (CDC) Atlanta, Georgia. This secondary analysis was not determined to meet the definition of "Research involving Human Subjects" and exempt from IRB review by Emory University (IRB00043231) on December 9, 2010.

Definitions

Genetic definitions: For all eight SNPs genotyped, samples were classified as (1) wild type or mutant and (2) pure or mixed, where a mixed sample contained PCR amplification of both wild type and mutant strains with the minor strain >30% of the major strain.

SP genotypes (*dhfr*, *dhps*, and combined *dhfr/dhps*) were determined according to the criteria outlined by Kublin and colleagues based on the number and type (mixed or

pure) of mutations in five SNPs: *dhfr-51,59,108* and *dhps-437,540* (16). Briefly, *dhfr* genotype, based on mutations in *dhfr*-51, 59, and 108, was classified as wild type, single, double, and triple (triple mixed and pure combined), *dhps* genotype (mutations in *dhps*-437 and 540) as wild type, single, and double (double mixed and pure combined), and combined *dhfr*/*dhps* genotype (mutations in *dhfr*-51,59,108 + *dhps*-437,540) as wild type, single, double, triple, quadruple, quintuple (quintuple mixed and pure combined). For SNPs linked to CQ resistance, *pfmdr1*-86 and *pfcrt*-76 were analyzed separately and genotypes were defined as wild type or mutant.

Mixed infections for all SNPs were considered mutant infections when calculating the prevalence of individual SNP mutations. The proportion of mixed infections was defined as the number of mixed infections divided by the total number of mutations (pure and mixed) in each SNP. Finally, *pfg377* microsatellite (MS) diversity, a genetic MS marker located within the coding region of the gametocyte maturation gene *pfg377*, was defined as infection with >1 allele. Appendix B provides a detailed description of laboratory procedures.

Clinical definitions: Epidemiological variables that were considered in the analysis included age, sex, parasite density, hemoglobin level (g/dL) as measured using the HemoCue system, presence of gametocytes, report of fever in previous 48 hours, geographic information system (GIS) distance (in meters) to the Lake Victoria shore, nearest clinic, nearest compound and elevation, and SP, CQ, and CTX drug use within two weeks prior to surveys. As CTX has anti-malarial properties and, similar to SP, acts on the parasite folic acid pathway, we assessed the relationship between antifolate drug

pressure and SP gene mutations by combining the usage of SP and/or CTX as a single variable we refer to as “antifolate”.

Data Analysis

Differences between participant characteristics at baseline and five years post-intervention were analyzed using chi-square for categorical variables and student’s t-tests (Satterthwaite’s statistic) for normally distributed continuous variables. Fisher’s exact tests were used when expected cell counts in contingency tables were less than five. Parasite density and GIS distance to nearest compound were log transformed prior to statistical testing, as the original variables were not normally distributed. Differences in the prevalence of SNP mutations, proportion of mixed infections in total mutations, and prevalence of SP genotypes between baseline and post-intervention samples were examined using chi-square tests.

To explore the association between epidemiological variables and drug resistance genotype, univariable and multivariable binary or cumulative logistic regression were used. SP resistance was studied considering mutations in *dhfr, dhps,* and *dhfr/dhps* combined as three separate outcomes. *Dhfr* mutations was analyzed using logistic regression by collapsing genotypes into two outcome categories in order to ensure sufficient sample size in each category for analysis: 1. wild type, single, and double collapsed, and 2. triple mutations. *Dhps* genotype was analyzed as three categories in the cumulative logistic model, without collapsing genotypes: 1. wild type, 2. single, and 3. double mutations. The combined *dhfr*/*dhps* genotype was classified into three categories to ensure adequate sample size in each group and analyzed using the cumulative logistic

regression: 1. wild type, single, double, and triple collapsed, 2. quadruple, and 3. quintuple mutations. For each of the SP resistance outcomes, two different multivariable models were used, one with antifolate use (SP and/or CTX) and another with only SP use as explanatory variables for drug use, due to expected multicollinearity if both variables were added to the same model. For each cumulative logistic model, the proportional odds assumption was evaluated using the score test, where $p<0.05$ reflected a violation of the assumption.

CQ resistance was studied by examining mutations in *pfcrt-76* and *pfmdr1-86* as separate outcomes. We also performed logistic regression to assess the association between epidemiological variables and mutations in *pfcrt*-76 or *pfmdr1*-86, considering wild type as the reference category. Appendix C provides a summary of the type of logistic regression model used for each of the outcomes analyzed.

For both binary and cumulative logistic regression methods, the final most parsimonious multivariable model was selected based biological plausibility, univariable analysis, and backwards elimination strategy (removal cutoff $p>0.10$) after assessing interaction and confounding. All possible two-way interaction terms with primary exposures of interest (drug use and survey year) and explanatory variables were examined using an overall likelihood ratio test. The likelihood ratio test was insignificant at the $\alpha = 0.05$ significance level for interaction terms in all models, so final statistical models were non-interaction models (Appendix F). Multicollinearity was assessed by examining condition indices and variance decomposition proportions (VDPs). Condition indices greater than 30 with corresponding VDPs larger than 0.5 were considered indicators of multicollinearity and, therefore, the variables elevation and log transformed

GIS distance to nearest compound were dropped from all models (Appendix G). We assessed confounding by assessing all possible combinations of remaining variables using a data-based criterion. In addition to the primary exposures of interest (drug use and survey year), the variables parasite density, age, sex, hemoglobin level, *pfg377* MS diversity, and GIS distance to shore were included in all models as potential confounders based on biologic rationale and previous studies. We compared the adjusted odds ratios (aOR) for the survey year variable in models containing all possible combinations of these explanatory variables with the remaining variables (fever, GIS distance to shore, and presence of gametocytes). If the aOR estimate was not within 10% of the gold standard model's aOR, the model was not considered to adequately control for confounding. The preferred model was selected to maximize precision from models that provided unconfounded aOR estimates (Appendix H). Finally, backwards elimination was used to remove variables that were not considered confounders from the model. The final model contained the variables drug use (SP, antifolate, or CQ) and survey year as the main predictors, and controlled for the potential confounders *pfg377* MS diversity, age, parasite density, hemoglobin level, sex, and GIS distance to shore.

We also assessed whether continuous variables were linearly associated with the log odds of the outcome, in order to use the continuous form of a variable in statistical models. Continuous variables were categorized and the adjusted odds ratios of the survey year variable were compared in models containing the continuous form of variables and categorical form of variables. As adjusted odds ratios were not meaningfully different, continuous variables were assumed to be linearly associated with the log odds of the outcome in univariable and multivariable regression models (Appendix I).

For all statistical tests, a two sided $p<0.05$ was considered to be statistically significant. Data analysis was performed using SAS software version 9.2 (SAS Institute Inc., Cary, NC, USA).

Results

Characteristics of Study Participants

Participants at the year 5 survey were significantly older, lived farther from the Lake Victoria shore and to the nearest clinic, and had a higher hemoglobin level, lower geometric mean parasite density, and increased proportion of *pfg377* MS diversity compared to those at baseline. Significantly more SP but less CQ use was reported by participants in the year 5 survey than in baseline. Year five survey participants were significantly more likely to report taking an antifolate (SP and/or CTX) within two weeks prior to the survey than baseline participants. Other characteristics including sex, report of fever, presence of gametocytes, elevation of compound, use of CTX, and use of any antimalarial did not differ significantly between year 5 and baseline survey samples (Table 1).

Prevalence of SNP Mutations and SP and CQ Genotypes

The prevalence of mutations and proportion of mixed infections by SNP at baseline and year 5 survey are summarized in Figure 3. The prevalence of SNP mutations in *dhfr* and in CQ resistance-linked genes was high initially and remained unchanged at the year 5 survey, with non-significant increases from 90.4% to 94.6% for *dhfr*-51, 74.8% to 81.8% for *dhfr-59,* 97.6% to 100% for *dhfr*-108, 81.6% to 81.8% for *pfcrt*-76, and a non-significant decrease from 75.1% to 73.4% for *pfmdr1*-86. In contrast, a statistically

significant increase ($p<0.001$) in the prevalence of mutations in both *dhps* codons was observed from baseline at 53.2% and 35.6% to the year 5 survey at 92.3% and 82.6% for *dhps*-437 and *dhps*-540, respectively (Figure 3a). At the year five survey, the proportion of mixed infections out of the total mutations (pure and mixed) was significantly lower at *dhfr*-51 ($p=0.0039$), *dhfr*-59 ($p=0.036$), *dhps*-437 ($p<0.0001$), and *dhps*-540 ($p<0.0001$) compared to baseline (Figure 3b). There was no significant change in the proportion of mixed infections measured in *pfcrt*-76 and *pfmdr1*-86 SNPs and all samples were pure mutant type for *dhfr*-108 except for one mixed sample at the year 5 survey. No mutations in *dhfr*-164 were detected at either survey.

The overall prevalence of mutations in *dhfr* ($p=0.0075$), *dhps* ($p<0.0001$), and *dhfr/dhps* combined ($p<0.0001$) genotypes was significantly different at baseline compared to the year 5 survey (Figure 4). The prevalence of the *dhfr* triple mutant increased from 66.7% to 76.5%, while the prevalence of the *dhps* double mutant changed more dramatically, from 29.6% to 78.5%, from baseline to the year 5 survey (Figure 4a, 4b). For *dhfr/dhps* combined, quintuple mutations increased from 29.3% at baseline to 62.0% at the year 5 survey corresponding to a decrease in double and triple mutations (Figure 4c).

Association Between Number of Mutations in SP and CQ Genotypes and Drug use or Year of Survey

The final model contained the variables drug use (SP, antifolate, or CQ) and year of survey as the main predictors, and controlled for age, parasite density, hemoglobin level, sex, *pfg377* MS diversity and GIS distance to shore as potential confounders. Other

variables considered but not included in the final model were presence of gametocytes, report of fever in previous 48 hours, presence of gametocytes and GIS distance to nearest clinic, nearest compound, and elevation, due to reasons described in the methods.

The relationship between drug use and drug resistance genotypes was studied by examining the effect of antifolate, SP, and CQ use on mutations in their corresponding molecular markers. Antifolate use was significantly associated with more mutations in the *dhps* genotype (adjusted odds ratio [OR], 2.4 [95% confidence interval {CI}, 1.2-5.1]) and, in the univariable model only, more mutations in *dhfr/dhps* combined genotype (unadjusted OR, 2.2 [95% CI, 1.2-3.8]) (Table 2). Among participants that reported using SP only, the unadjusted odds for having more *dhps* mutations (unadjusted OR, 5.6 [95% CI, 1.6-19.4]) and more *dhfr/dhps* combined mutations (unadjusted OR, 3.0 [95% CI, 1.2-7.3]) were significantly higher than for those who reported no SP use (Table 2). These associations did not remain significant after adjusting for sex, age, hemoglobin level, parasite density, *pfg377* MS diversity, and GIS distance to shore as well as year of survey. No significant associations between CQ use and mutations in *pfcrt*-76 or *pfmdr1*-86 were observed in either univariable or multivariable analysis (Table 3).

The associations between the year of survey and SP and CQ genotypes were also explored. At the year 5 survey, the unadjusted odds of having more mutations in *dhfr*, *dhps*, and *dhfr/dhps* combined genotypes, respectively, were 1.6 ([95% CI, 1.1-1.4]), 9.2 ([95% CI, 6.2-13.7]), and 7.5 ([95% CI, 5.2-10.8]) times the odds of having more mutations at baseline (Table 2). In multivariable analysis, the association remained significant for *dhps* genotype (adjusted OR, 10.3 [95% CI, 6.2-17.2]) and *dhfr/dhps* combined genotype (adjusted OR, 8.5 [95% CI, 5.3-13.6]) (Table 2) after adjusting for

antifolate drug use and other potential confounders. The year of survey was not associated with the *pfcrt*-76 or *pfmdr1*-86 mutant in either univariable or multivariable analysis (Table 3).

Discussion

Until an effective vaccine becomes available, malaria prevention and treatment will rely on vector control tools and the use of antimalarial drugs. As the development of new drugs is expensive and the implementation of potentially ineffective partner drugs can pose a treatment risk, preserving the efficacy of currently available antimalarial drugs by monitoring and limiting the spread of drug resistance is imperative (7). With the scale-up of ITNs as a malaria control strategy (9), assessing the impact of the transmission-reducing intervention on the spread of drug resistance has become particularly important. The current understanding of the relationship between transmission intensity and the spread of drug resistance relies on the effect of host, parasite, and vector factors (8, 12). In this study, we investigated the effect of sustained transmission reduction by ITNs on the spread of drug resistance during a shift in national drug policy from CQ to SP. We compared the prevalence of gene mutations associated with drug resistance to SP and CQ at baseline and five years post-ITN introduction and further explored how some clinical, epidemiological, and genetic variables shape the dynamics of drug resistance.

The accumulation of mutations in the *dhfr* and *dhps* genes in response to drug pressure is stepwise and occurs initially in the *dhfr* gene. We observed a high prevalence of the *dhfr* triple mutant parasite (67%), but a relatively lower prevalence of the *dhps* double mutant (30%) at baseline. Although SP use in the study area was limited due to

poor availability prior to the national drug policy change in 1998 (25), the high prevalence of mutations in *dhfr* gene at the baseline (1996) could be attributed to both the on-going use of CTX, which is also an inhibitor of the *dhfr* and *dhps* enzymes (15), and the possible gene flow due to human migration from the surrounding bay area, where SP use was relatively common (26) , to our study area. At the year 5 survey, there was a significant increase in the prevalence of the *dhps* double mutant (from 30% to 79%) and *dhfr/dhps* quintuple mutations (from 20% to 62%), but a less dramatic increase in *dhfr triple* mutant (from 67% to 77%) (Figure 4). Statistical modeling showed that SP use alone and antifolate use (SP and/or CTX) were associated with more mutations in *dhps* genotype and *dhfr/dhps* combined genotype, although only the association between antifolate use with more mutations in *dhps* genotype remained statistically significant in multivariate regression (Table 2). Given these findings, it is clear that increased antifolate drug use, most likely resulting from drug policy change, played a predominant role in the selection of SP resistant parasites, leading to the high prevalence of *dhps* double and *dhfr/dhps* quintuple mutations during the period of ITN intervention. Our results from the children in this study are consistent with the increased prevalence of *dhfr* and *dhps* mutations observed in all age groups in Kisumu, Kenya as well as in Tanzania after drug policy changed to SP or CTX (27-29).

We further observed a strong and significant association between the year 5 survey and more mutations in *dhps* and *dhfr/dhps* combined genotypes (Table 2). This association remained distinct after adjusting for antifolate drug use and potential confounders, suggesting that the differences in factors between the two survey time points not directly measured in this study may have regulated the high prevalence of SP

mutations observed at the year 5 survey. Our speculation could rely on the following plausible explanations of a few proxy factors.

Reduction in transmission intensity presumably lowers the number of malaria infections per human host (clonal multiplicity) (8). However, a previous investigation on a subset of the samples used in the present study demonstrated an unchanged, high overall clonal multiplicity despite the reduction in EIR five years post-ITN intervention, suggesting strong resiliency of the malaria parasite in response to dramatic transmission reduction after five years of sustained ITN use (30). Although the unchanged high level of clonal multiplicity alone has no effect on the spread of drug resistance for the monogenic-based SP drug resistant genes (*dhfr* and *dhps*), the rate of drug resistance could increase in the presence of intrahost competition between co-infecting parasite clones (intrahost dynamics) based on the generalized immunity model (13). In the current study, we observed a significant decrease in the proportion of mixed infections out of total mutations for *dhfr*-51,59 and *dhps*-437,540 SNPs at the year 5 survey compared to baseline (Figure 3b). This result suggests that intrahost removal of SP drug wild/sensitive parasite clones is present at the year 5 survey, thus selecting for and expanding pure drug resistance at the population level. However, the exact effect of altered host immunity due to the transmission reduction by ITN use (not measured here) with increased antifolate drug use on the removal of SP drug sensitive parasites remains unclear (13).

At baseline, the prevalence of *pfcrt*-76 and *pfmdr1*-86 mutations was high at 82% and 75%, respectively, which presumably resulted from the progression of CQ drug pressure prior to policy change. Despite decreased CQ drug use due to the change in first line antimalarial treatment to SP during the second year of the ITN trial and decreased

malaria case treatment due to the transmission reduction by use of ITNs (25), the high prevalence of CQ-resistant gene mutations remained unchanged between baseline and the year 5 survey (82% and 73%, respectively). This result could have several explanations. First, the time frame evaluated in this study may not be sufficient to observe considerable decreases in mutation prevalence of CQ molecular markers at such high levels; therefore, long term monitoring is necessary. A study conducted in Malawi showed that the *pfcrt* mutant *g*enotype significantly declined after cessation of CQ use for eight years (31). Second, CQ was not completely withdrawn from Kenya after the drug policy shift to SP in 1998 as shown in Table 1, hence, CQ drug pressure may not have dramatically decreased. Third, the unchanged high clonal multiplicity measured by a previous investigation on a subset of this study's samples (30) and the absence of intrahost competition measured by the unchanged proportion of mixed infections in CQ drug resistant markers in the current study suggest that transmission reduction by use of ITNs does not affect the prevalence of CQ mutations. This explanation supports the conceptual frameworks described earlier (8, 12) .

Strengths and Limitations of Study

Many published studies on antimalarial drug resistance have presented molecular data and studied the association between molecular drug resistance outcomes and epidemiological variables, but few studies have considered the effect of transmission reduction by vector control on antimalarial drug resistance. By linking clinical and epidemiological characteristics with molecular marker results, we were able to investigate the effect of survey year (before or 5 years post-ITN intervention) and drug use on antimalarial drug resistance genes during a period of drug policy change. We

compared samples prior to and five years post-ITN introduction, which allowed for the evaluation of sustained effects rather than short-term, possibly bottleneck effects, on the prevalence of parasite drug resistance genes. Finally, although our study was conducted in children under the age of five, globally, this population bears the brunt of the malaria disease burden; therefore, our findings are relevant in a population that is most vulnerable to malaria.

There were a few limitations of this study. The question posed in this research was not a primary research question for the ITN trial and this limitation had several important implications in the study design and interpretation of the results. First, formal statistical sample size calculations were not performed. However, the sample size used in this study is consistent with similar previously published studies on antimalarial drug-resistance. Second, the ITN trial did not include a nearby comparison area without community ITN use at the year 5 survey, which limited comparisons to dissect and quantify the contribution of transmission reduction by ITN use on the spread of drug resistance at the same time point. Third, it is not clear whether drug use as measured in the current study reflects the proportion of therapeutic drug use that is mainly influenced by acquired immunity and/or the level of community drug use that is regulated by infection risk. Consequently, we were unable completely to explain the role of drug pressure in our results. It would be ideal to conduct well-controlled studies with comparable sites where ITNs have not been distributed to date or in the areas with dramatically different levels of ITN coverage/usage during the same time period with no drug policy change. Such studies would help to assess the net effect of transmission reduction by ITNs on the spread of antimalarial drug resistance. Finally, the results of this

study may not be generalizable in areas where the epidemiology of malaria and coverage of malaria control interventions are different.

Current Study in the Context of Previously Published Findings

Our findings differ with the results from the studies conducted in Tanzania and Burkina Faso which also assessed the short or long term effects of ITNs or ITCs on prevalence of gene mutations linked to SP and CQ, respectively (18, 20). The discrepancy in the results among the studies could be due to differences in the study design, level of transmission reduction, stage of existing drug resistance, change in drug policy during the study, and other unmeasured potential confounders. In the Tanzania study, a reduction in the prevalence of *dhfr* triple mutation was observed two years after ITN introduction and during the two year study period SP was first line drug treatment for malaria. The study conducted in Burkina Faso, where CQ remained first line treatment during study period, reported no change in the prevalence of molecular markers linked to CQ and SP after seven years of ITC intervention. Our study showed that increased antifolate drug use likely led to the high prevalence of SP mutations five years post-ITN intervention and reduced transmission did not change the existing high prevalence of CQ mutations. In addition, the difference in the degree of local gene flow resulting from human movement between intervention and non-intervention areas could be another factor for the inconsistent results among the different studies. Although the previous short-term study concluded that transmission reducing interventions such as ITNs may help restore susceptibility to SP (18), there is no evidence from the current study that sustained transmission reduction by ITNs reduces the prevalence of drug resistance genes associated with SP and CQ.

REFERENCES

1. World Health Organization. World malaria report 2010. Geneva: World Health Organization; 2010.

2. Malaria: About Malaria. In: Centers for Disease Control and Prevention.

3. Sachs J, Malaney P. The economic and social burden of malaria. Nature 2002;415(6872):680-5.

4. Kenya Malaria Fact Sheet. In: Malaria Control Program: Kenya; 2011.

5. D'Alessandro U, Buttiens H. History and importance of antimalarial drug resistance. Trop Med Int Health 2001;6(11):845-8.

6. Amin AA, Zurovac D, Kangwana BB, Greenfield J, Otieno DN, Akhwale WS, et al. The challenges of changing national malaria drug policy to artemisinin-based combinations in Kenya. Malar J 2007;6:72.

7. Zimmerman PA. Roll back of Plasmodium falciparum antifolate resistance by insecticide-treated nets. Am J Trop Med Hyg 2003;69(3):236-7.

8. Hastings IM, Watkins WM. Intensity of malaria transmission and the evolution of drug resistance. Acta Trop 2005;94(3):218-29.

9. Lengeler C. Insecticide-treated nets for malaria control: real gains. Bull World Health Organ 2004;82(2):84.

10. Gamble C, Ekwaru JP, ter Kuile FO. Insecticide-treated nets for preventing malaria in pregnancy. Cochrane Database Syst Rev 2006(2):CD003755.

11. Gimnig JE, Vulule JM, Lo TQ, Kamau L, Kolczak MS, Phillips-Howard PA, et al. Impact of permethrin-treated bed nets on entomologic indices in an area of intense year-round malaria transmission. Am J Trop Med Hyg 2003;68(4 Suppl):16-22.

12. Talisuna AO, Okello PE, Erhart A, Coosemans M, D'Alessandro U. Intensity of malaria transmission and the spread of Plasmodium falciparum resistant malaria: a review of epidemiologic field evidence. Am J Trop Med Hyg 2007;77(6 Suppl):170-80.

13. Hastings IM, D'Alessandro U. Modelling a predictable disaster: the rise and spread of drug-resistantmalaria. Parasitol Today 2000;16(8):340-7.

14. Olliaro P. Mode of action and mechanisms of resistance for antimalarial drugs. Pharmacol Ther 2001;89(2):207-19.

15. Iyer JK, Milhous WK, Cortese JF, Kublin JG, Plowe CV. Plasmodium falciparum cross-resistance between trimethoprim and pyrimethamine. Lancet 2001;358(9287):1066-7.

16. Kublin JG, Dzinjalamala FK, Kamwendo DD, Malkin EM, Cortese JF, Martino LM, et al. Molecular markers for failure of sulfadoxine-pyrimethamine and chlorproguanil-dapsone treatment of Plasmodium falciparum malaria. J Infect Dis 2002;185(3):380-8.

17. Wongsrichanalai C, Pickard AL, Wernsdorfer WH, Meshnick SR. Epidemiology of drug-resistant malaria. Lancet Infect Dis 2002;2(4):209-18.

18. Alifrangis M, Lemnge MM, Ronn AM, Segeja MD, Magesa SM, Khalil IF, et al. Increasing prevalence of wildtypes in the dihydrofolate reductase gene of Plasmodium falciparum in an area with high levels of sulfadoxine/pyrimethamine resistance after introduction of treated bed nets. Am J Trop Med Hyg 2003;69(3):238-43.

19. Mharakurwa S, Mutambu SL, Mudyiradima R, Chimbadzwa T, Chandiwana SK, Day KP. Association of house spraying with suppressed levels of drug resistance in Zimbabwe. Malar J 2004;3:35.

20. Diallo DA, Sutherland C, Nebie I, Konate AT, Ord R, Pota H, et al. Sustained use of insecticide-treated curtains is not associated with greater circulation of drug-resistant malaria parasites, or with higher risk of treatment failure among children with uncomplicated malaria in Burkina Faso. Am J Trop Med Hyg 2007;76(2):237-44.

21. Lindblade KA, Eisele TP, Gimnig JE, Alaii JA, Odhiambo F, ter Kuile FO, et al. Sustainability of reductions in malaria transmission and infant mortality in western Kenya with use of insecticide-treated bednets: 4 to 6 years of follow-up. JAMA 2004;291(21):2571-80.

22. Phillips-Howard PA, ter Kuile FO, Nahlen BL, Alaii JA, Gimnig JE, Kolczak MS, et al. The efficacy of permethrin-treated bed nets on child mortality and morbidity in western Kenya II. Study design and methods. Am J Trop Med Hyg 2003;68(4 Suppl):10-5.

23. Singer LM, Mirel LB, ter Kuile FO, Branch OH, Vulule JM, Kolczak MS, et al. The effects of varying exposure to malaria transmission on development of antimalarial antibody responses in preschool children. XVI. Asembo Bay Cohort Project. J Infect Dis 2003;187(11):1756-64.

24. ter Kuile FO, Terlouw DJ, Kariuki SK, Phillips-Howard PA, Mirel LB, Hawley WA, et al. Impact of permethrin-treated bed nets on malaria, anemia, and growth in infants in an area of intense perennial malaria transmission in western Kenya. Am J Trop Med Hyg 2003;68(4 Suppl):68-77.

25. Phillips-Howard PA, Wannemuehler KA, ter Kuile FO, Hawley WA, Kolczak MS, Odhacha A, et al. Diagnostic and prescribing practices in peripheral health facilities in rural western Kenya. Am J Trop Med Hyg 2003;68(4 Suppl):44-9.

26. Terlouw DJ, Nahlen BL, Courval JM, Kariuki SK, Rosenberg OS, Oloo AJ, et al. Sulfadoxine-pyrimethamine in treatment of malaria in Western Kenya: increasing resistance and underdosing. Antimicrob Agents Chemother 2003;47(9):2929-32.

27. Spalding MD, Eyase FL, Akala HM, Bedno SA, Prigge ST, Coldren RL, et al. Increased prevalence of the pfdhfr/phdhps quintuple mutant and rapid emergence of pfdhps resistance mutations at codons 581 and 613 in Kisumu, Kenya. Malar J 2010;9:338.

28. Hamel MJ, Greene C, Chiller T, Ouma P, Polyak C, Otieno K, et al. Does cotrimoxazole prophylaxis for the prevention of HIV-associated opportunistic infections select for resistant pathogens in Kenyan adults? Am J Trop Med Hyg 2008;79(3):320-30.

29. Malisa AL, Pearce RJ, Abdulla S, Mshinda H, Kachur PS, Bloland P, et al. Drug coverage in treatment of malaria and the consequences for resistance evolution--evidence from the use of sulphadoxine/pyrimethamine. Malar J 2010;9:190.

30. Gatei W, Kariuki S, Hawley W, ter Kuile F, Terlouw D, Phillips-Howard P, et al. Effects of transmission reduction by insecticide-treated bed nets (ITNs) on parasite genetics population structure: I. The genetic diversity of Plasmodium falciparum parasites by microsatellite markers in western Kenya. Malar J 2010;9:353.

31. Laufer MK, Thesing PC, Eddington ND, Masonga R, Dzinjalamala FK, Takala SL, et al. Return of chloroquine antimalarial efficacy in Malawi. N Engl J Med 2006;355(19):1959-66.

32. Lindblade K. Personal communication. In; 2011.

33. Alker AP, Mwapasa V, Meshnick SR. Rapid real-time PCR genotyping of mutations associated with sulfadoxine-pyrimethamine resistance in Plasmodium falciparum. Antimicrob Agents Chemother 2004;48(8):2924-9.

34. Purfield A, Nelson A, Laoboonchai A, Congpuong K, McDaniel P, Miller RS, et al. A new method for detection of pfmdr1 mutations in Plasmodium falciparum DNA using real-time PCR. Malar J 2004;3:9.

35. Wilson PE, Kazadi W, Kamwendo DD, Mwapasa V, Purfield A, Meshnick SR. Prevalence of pfcrt mutations in Congolese and Malawian Plasmodium falciparum isolates as determined by a new Taqman assay. Acta Trop 2005;93(1):97-106.

36. Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, et al. Microsatellite markers reveal a spectrum of population structures in the malaria parasite Plasmodium falciparum. Mol Biol Evol 2000;17(10):1467-82.

TABLES

Table 1. Characteristics of study participants at baseline (1996) and year 5 survey (2001) among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya

Category/Characteristics	Baseline (n=250)	Year 5 Survey (n=242)	p-value
Male sex (%)	126/250 (50.4)	121/242 (50.0)	0.93
Age, mean (SD), months	23.8 (15.0)	43.3 (17.6)	<0.0001*
Fever[a] (%)	23/239 (9.6)	21/241 (8.7)	0.73
Hemoglobin, mean (SD), g/dL	7.9 (2.2)	10.2 (1.7)	<0.0001*
Parasite density[b], geometric mean (SD), uL	2670 (5933)	1339 (9582)	<0.0001*
Gametocytes present (%)	43/250 (17.2)	56/242 (23.1)	0.1
Pfg377 MS diversity[c] (%)	102/246 (43.2)	129/237 (54.4)	0.015*
Euclidean distance, mean (SD), meters			
To shore	7029 (2657)	7582 (2530)	0.018*
To nearest clinic	2255 (1017)	2534 (970)	0.0019*
Elevation of compound	1247 (52)	1251 (46)	0.33
Drug use[d] (%)			
SP	4/246 (1.6)	17/236 (7.2)	0.0027*
CTX	13/247 (5.3)	20/242 (8.3)	0.19
SP and/or CTX (antifolate)	16/250 (6.4)	36/242 (14.9)	0.0022*
CQ	46/247 (18.6)	25/236 (10.6)	0.013*
CQ or SP	50/250 (20.0)	42/242 (17.4)	0.45
Bednet usage[e] (%)	--	223/242 (92.2)	--

NOTE. Data are proportion (%) of *P. falciparum* smear-positive participants with molecular data, unless otherwise indicated.

Abbreviations: SD, standard deviation; MS, microsatellite; SP, sulfadoxine-pyrimethamine; CTX, cotrimoxazole; CQ, chloroquine.

[a] Body temperature greater than or equal to 37.5°C at or 48 hours before survey.

[b] Parasite density was log transformed prior to statistical analysis.

[c] Infection with more than one allele based on microsatellite marker in coding region of *pfg377*.

[d] Within two weeks prior to survey.

[e] Child slept under bednet >5 days in the past week, bednet usage in the study area was <5% prior to trial.

* P<0.05, statistically significant difference between baseline and year 5 survey based on chi-squared or two sample t-test (satterthwaite).

Table 2. Univariable and multivariable analyses of the association between specific predictors and mutations in *dhfr*, *dhps*, and *dhfr/dhps* combined among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya

	dhfr Triple Mutant[a]		*dhps* Mutations[b]		*dhfr/dhps* Mutations[c]	
	OR (95% CI)		OR (95% CI)		OR (95% CI)	
Predictor	Unadjusted	Adjusted[d]	Unadjusted	Adjusted[e]	Unadjusted	Adjusted[e]
Drug use						
SP and/or CTX (antifolate)[f]	1.1 (0.6, 2.1)	1.0 (0.5, 2.1)	3.2 (1.6, 6.2)*	2.4 (1.2, 5.1)*	2.2 (1.2, 3.8)*	1.7 (0.9, 3.2)
SP only[g]	1.0 (0.4, 2.7)	0.9 (0.3, 2.4)	5.6 (1.6, 19.4)*	3.2 (0.9, 12.0)	3.0 (1.2, 7.3)*	1.9 (0.7,5.0)
Year of Survey[h]	1.6 (1.1, 2.4)*	1.5 (0.9, 2.5)	9.2 (6.2, 13.7)*	10.3 (6.2, 17.2)*	7.5 (5.2, 10.8)*	8.5 (5.3, 13.6)*

Abbreviations: CI, confidence interval; OR, odds ratio; SP, sulfadoxine-pyrimethamine; CTX, cotrimoxazole.

[a] *dhfr* Triple Mutant was analyzed by grouping wild type, single and double genotypes as the reference category.

[b] *dhps* mutations were analyzed as 3 genotype categories: (1) wild type, (2) single, and (3) double.

[c] *dhfr* and *dhps* combined mutations were analyzed as 3 genotype categories: (1) wild type, single, double, and triple, (2) quadruple, and (3) quintuple.

[d] Derived from multivariable logistic regression, controlling for age, sex, parasite density, hemoglobin level, *pfg377* microsatellite diversity and GIS distance to shore.

[e] Derived from multivariable cumulative logistic regression, which models the probability of more mutations compared to fewer, controlling for age, sex, parasite density, hemoglobin level, *pfg377* microsatellite diversity and GIS distance to shore.

[f] Model containing SP and/or CTX use as a predictor (antifolate), N=460.

[g] Model containing SP use only as a predictor, N=450.

[h] Year 5 survey compared to baseline.

* Statistically significant, $p<0.05$.

Table 3. Univariable and multivariable analyses of the association between specific predictors and mutations in CQ-linked drug resistance genes among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya

	pfcrt-76 Mutant[a] (N=452)		*pfmdr1*-86 Mutant[a] (N=451)	
	OR (95% CI)		OR (95% CI)	
Predictor	Unadjusted	Adjusted[b]	Unadjusted	Adjusted[b]
CQ use	1.0 (0.5, 1.9)	1.2 (0.6, 2.4)	1.5 (0.8, 2.7)	1.7 (0.9, 3.3)
Year of Survey[c]	1.0 (0.6, 1.6)	1.3 (0.7, 2.3)	0.9 (0.6, 1.4)	0.8 (0.5, 1.3)

Abbreviations: CI, confidence interval; OR, odds ratio; CQ, chloroquine.
[a] *pfcrt*-76 and *pfmdr1*-86 mutants were analyzed using wild type as the reference group.
[b] Derived from multivariable logistic regression, controlling for age, sex, parasite density, hemoglobin level, *pfg377* microsatellite diversity and GIS distance to shore.
[c] Year 5 survey versus baseline.

FIGURES/FIGURE LEGENDS

Figure 1. Conceptual framework for relationship between transmission intensity and anti-malarial drug resistance (adapted from Hastings IM and Watkins WM, 2005 (8))

Figure 2. Flow Diagram of ITN trial and drug resistance study among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya (21, 22, 24, 32)

Figure 3. Comparison of mutation prevalence by SNP between baseline (1996) and year 5 survey (2001) among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya. *A,* Overall prevalence of mutations ((pure mutations and mixed) / total samples). *B,* Proportion of mixed infections in total mutations (mixed / (pure mutations and mixed)). Statistical analysis performed using chi-squared test. * $p<0.05$, significant difference in prevalence between baseline and year 5 surveys.

Figure 4. Prevalence of SP genotypes at baseline (1996) and year 5 survey (2001) among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya. *A, dhfr* genotype based on mutations in *dhfr*-51,59,108. *B, dhps* genotype based on mutations in *dhps*-437,540. *C, dhfr/dhps* combined genotype based on *dhfr* and *dhps* genotypes. Statistical analysis performed using chi-squared test. The prevalence of *dhfr*, *dhps*, and *dhfr/dhps* combined genotypes were significantly different at baseline compared to the year 5 survey, $p<0.05$.

Figure 1. Conceptual framework for relationship between transmission intensity and anti-malarial drug resistance

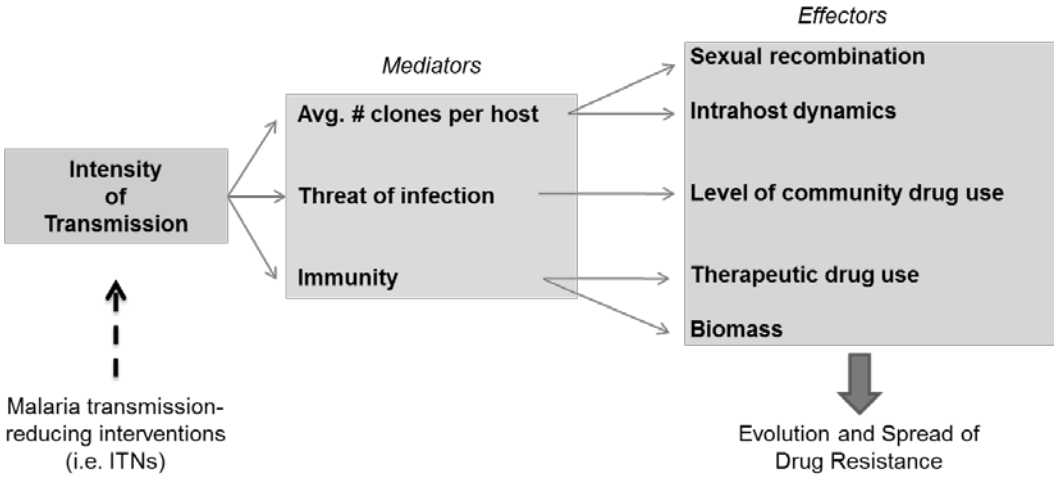


Figure 2. Flow Diagram of western Kenya ITN trial and drug resistance study samples collected from *P. falciparum* smear-positive children under the age of five

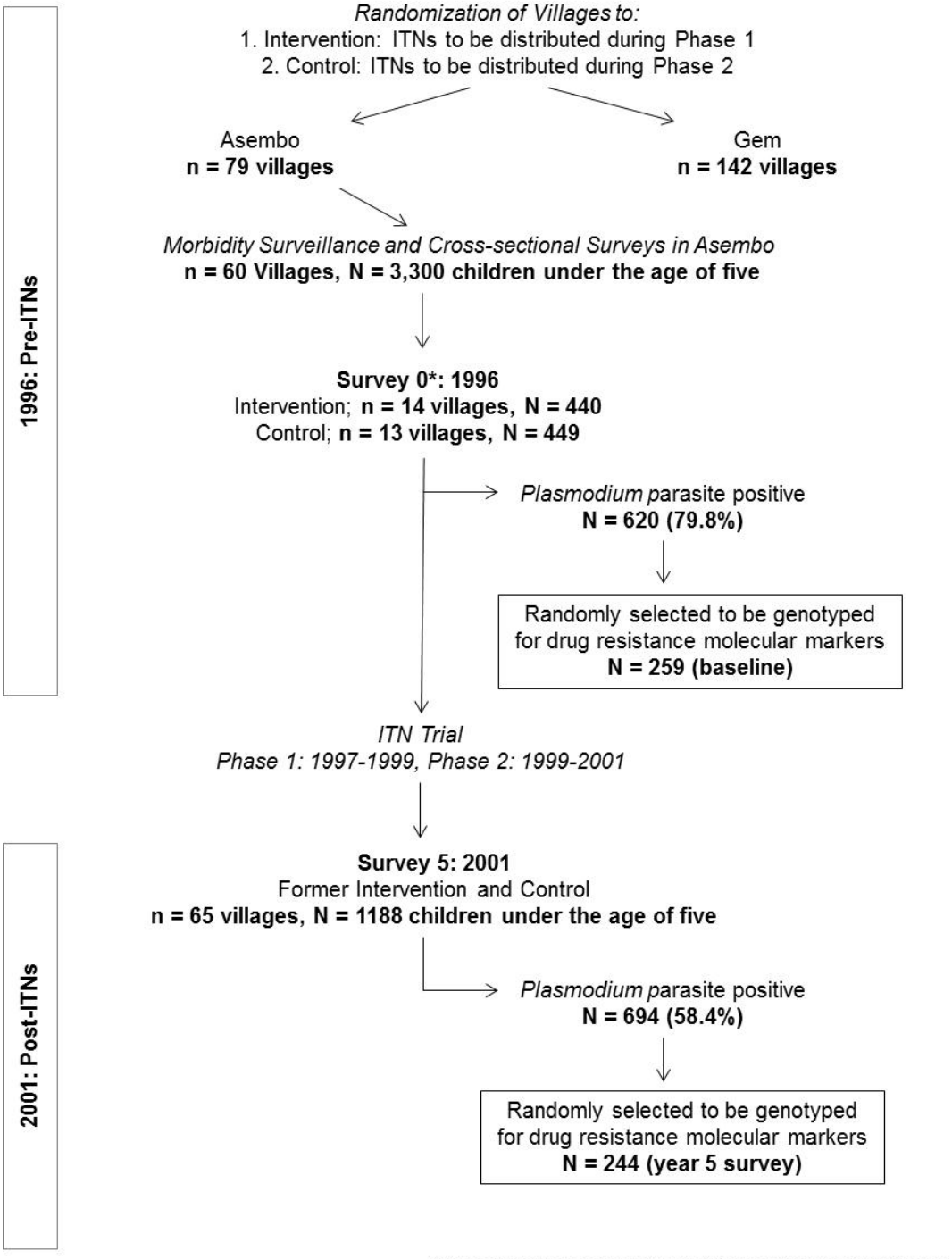


Figure 3. Comparison of mutation prevalence by SNP between baseline (1996) and year 5 survey (2001) among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya

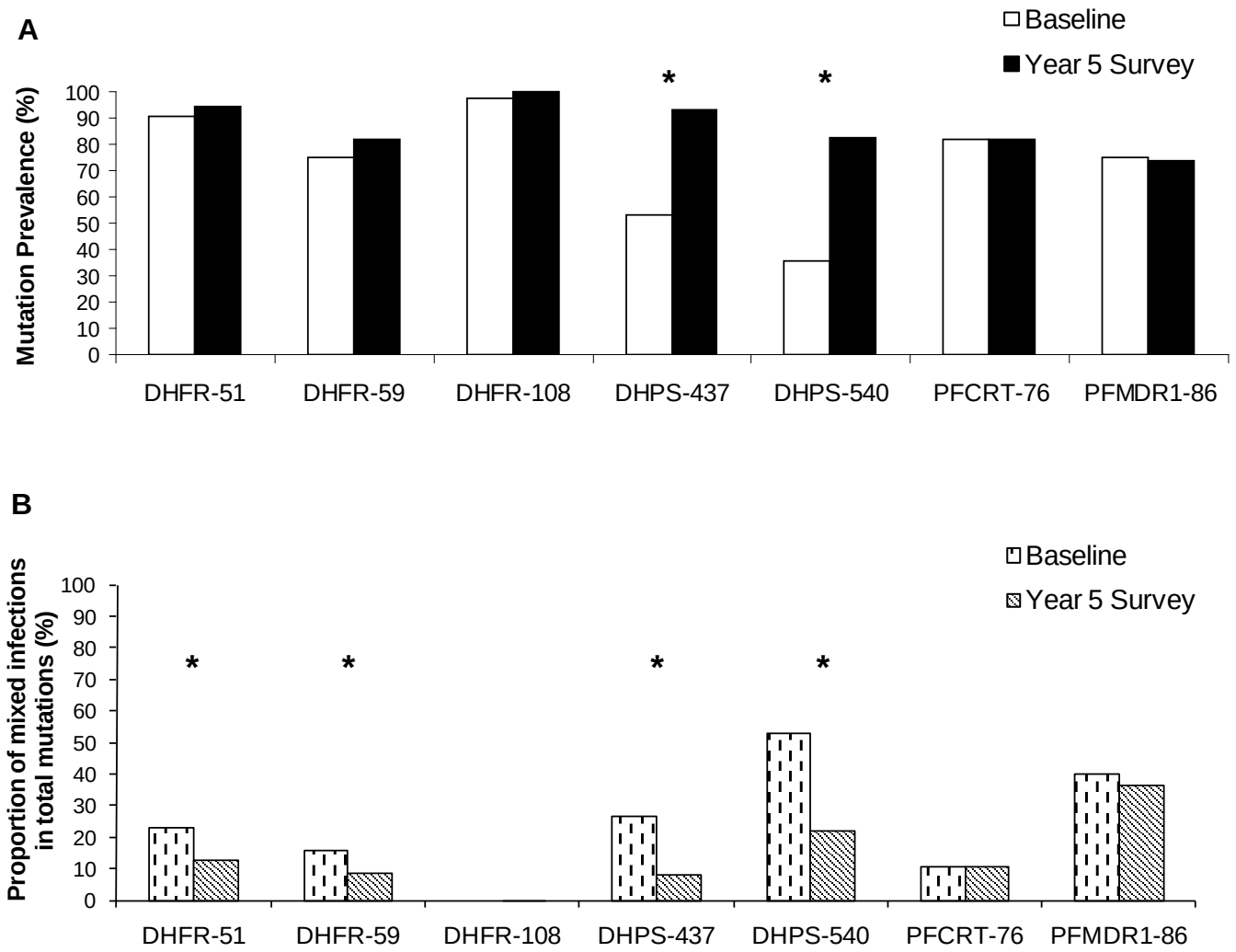
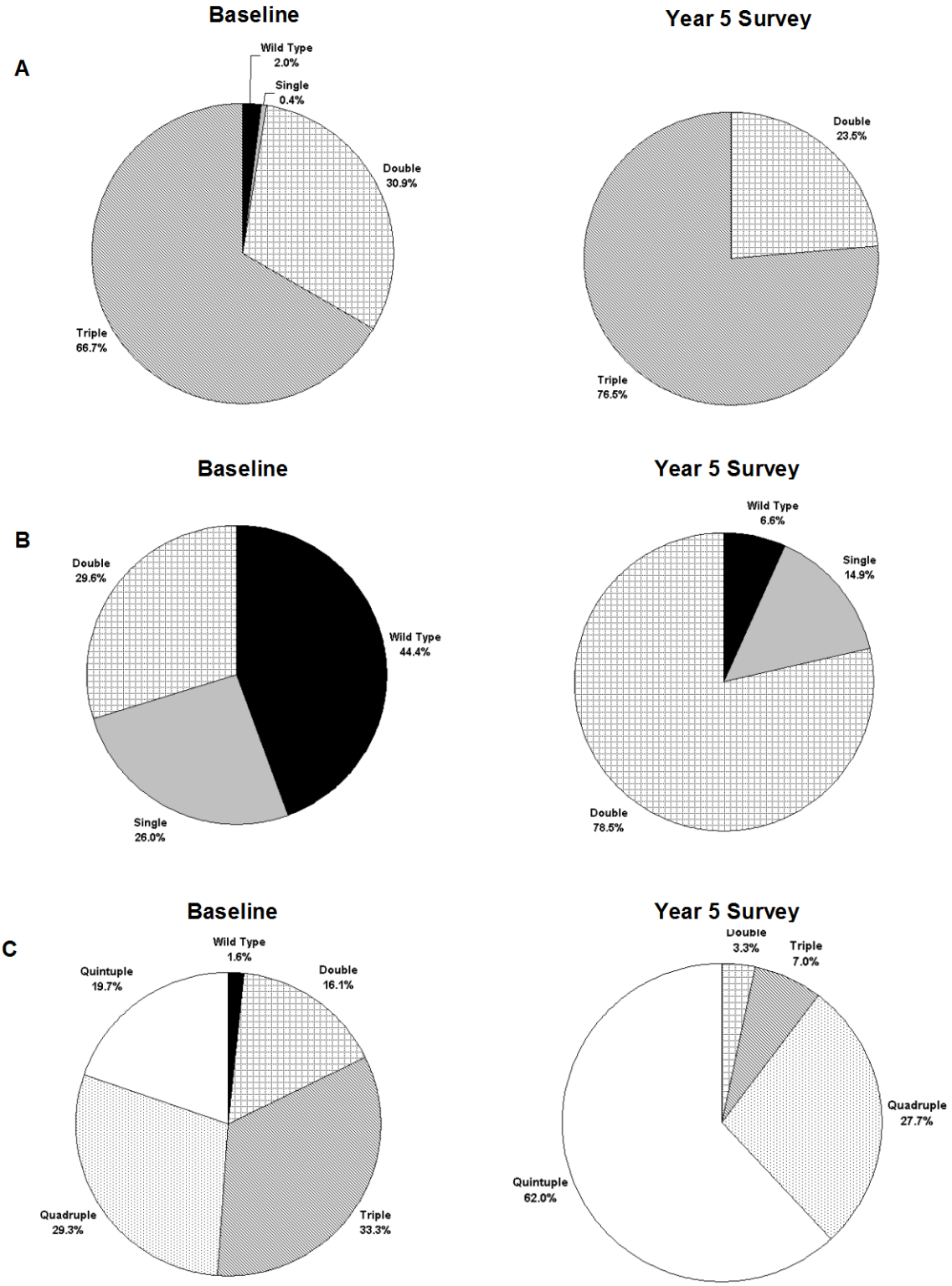


Figure 4. Prevalence of SP genotypes at baseline (1996) and year 5 survey (2001) among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya



CHAPTER III: SUMMARY, PUBLIC HEALTH IMPLICATIONS, POSSIBLE FUTURE DIRECTIONS

Despite the introduction of wide-spread implementation of malaria prevention tools, improved diagnostics, and effective ACT treatment, malaria remains a global public health challenge. Significant reductions in malaria burden must be achieved for many countries to meet Millennium Development Goal (MDG) 2015 target 4, reduce by two-thirds the mortality rate among children under five, and target 6, begin to reverse the incidence of malaria and other major diseases (1). The emergence and spread of drug resistance to antimalarial drugs could substantially limit the progress towards MDGs and reducing and burden of malaria world-wide. As malaria prevention and treatment strategies are implemented concurrently in individuals and populations, any potential interaction—whether beneficial or harmful—between interventions can have important implications on malaria control (7). If transmission-reducing interventions facilitate the spread of drug resistance, antimalarial drug policy may need to be changed sooner than anticipated. However, if transmission-reducing interventions decrease the spread of drug resistance, currently used antimalarial drugs may remain efficacious for a longer period of time, thus decreasing expenditure on drug development and policy implementation and saving more lives by reducing treatment failure (7, 8, 12).

We investigated whether transmission reduction by ITNs affects the prevalence of drug resistant genes associated with SP and CQ in the *P. falciparum* parasite during a five-year period in western Kenya when antimalarial treatment policy changed from CQ to SP. Our findings suggest that increased antifolate drug use most likely due to policy change was associated with an increased prevalence of SP drug resistant genes, while the

high prevalence of CQ resistance genes remained unchanged. There is no evidence from the current study that substantial transmission reduction of malaria by ITNs had a measurable effect on drug resistance. Our study has important implications in molecular surveillance of antimalarial drug resistance in the areas with high coverage of ITNs. Antimalarial drug resistance should continue to be monitored by molecular marker surveillance in order to understand whether transmission-reducing interventions affect the prevalence of drug resistance.

APPENDICES

Appendix A. Study questionnaires

Cross-section survey questionnaires 0 (BX0, baseline) and 5 (BX5, year 5 survey) were similar to the questionnaire provided below for survey 4 (BX4):

Filing Number	~~~~	*Bednet X4 Code (Team # Village # / child #)*	*PLACE LABEL HERE*

Bednet Cross-Sectional Survey X4 May-June 2000
Registration and Clinical Form

Part 1: Registration

1.01	Today's date *(day/month/year)*	~~/~~/~~	1.02 Villlage/Comp, House	~~/~~~,~			
1.03	Relationship of person accompanying the child today *(1= mother, 2=father, 3= grandparent, 4=aunt, 5=co-wife, 6= sibling, 7= other)*	~	1.04 Is the birth mother of this child alive now? *Y/N*	~			
1.05	If no, did she die within the last year? *Y/N*	~	1.06 Is the father of this child alive now? *Y/N/D/U*	~	1.07 If no, did he die within the last year? *Y/N*	~	
1.08	Mother's name *(natural mother)*	*Christian name*	1.09 *Juok name*	1.10 *Husbands juok name*			
1.11	Mother's clan *(husband's clan if married)*	*Clan name*	1.12 Mother's date of birth	~~/~~/~~	1.13 Mother's age *(years)*	~~	
1.14	Child's name	*Christian name*	1.15 *Juok name*	1.16 *Father's juok name*			
1.17	Child's DOB *(If not known, copy from the list)*	~~/~~/~~	1.18/1.19 Child's age	*years* ~ *months* ~~	1.20 Sex of child *(M/F)*	~	
1.21	Where does child normally sleep? *(1=bedroom 2=kitchen 3= sitting room 4=other 8=DK)*	~	1.22 Do you have a net in the house? *(If yes go to 1.23, if no go to 2.01)*	*Y/N/D* ~			
1.23	If yes; Λ is it the green bednet supplied by CDC/KEMRI project? *Y/N/D*	~	1.24 Is the child sleeping under this net more than 5 days per week? *Y/N/D*	~			

Part 2: Vaccinations

If child is 24 months or older, Λ go to 3.01			
2.01	Has this child been vaccinated against childhood diseases? *Bende nyathi oseyudo chanjo?*	*Y/N/D* ~	
2.02	Where did the child go for his/her last vaccination? *(1=Abidha 2=Ong'ielo 3=Lwak 4=Saradidi 5=Ndori 6=Gobei 7=Nyagoko 10=Rarieda 11=Masala 12=Komolo 13=Bondo 14=Rakoyo 15=Mahaya 16=Other 17=NA)*	~~	
2.03	Do you have the vaccination card/booklet?	*Y/N* ~	
If yes; use card to copy the vaccination status (Y/N/D) and dates. If no; ask for each of the vaccination types in questions 2.04 to 2.09 if the child was vaccinated for this type (Y/N/D) and ask the approximate dates of vaccination		***Y/N/D***	***Date*** *(day/month/year) If the exact dates are not known fill in the year and if possible the month*
2.04	BCG vaccine (at birth)	~	~~/~~/~~
2.05	Polio vaccine (at birth)	~	~~/~~/~~
2.06	DPT 1 / OPV1 (Polio) (at around 6 weeks)	~	~~/~~/~~
2.07	DPT 2 / OPV2 (Polio) (at around 10 weeks)	~	~~/~~/~~
2.08	DPT 3 / OPV3 (Polio) (at around 14 weeks)	~	~~/~~/~~
2.09	Measles (at around 9 months)	~	~~/~~/~~

Part 3: Symptoms of illness of survey child reported by caretaker

3.01	Is this child ill now? ***(If yes: Λ continue with 3.02. If no:Λ go to 3.03)***	***Y/N/D*** ~	3.02	Does this child have a fever now? ***(Let the caretaker feel the child)***	***Y/N/D*** ~
3.03	Including today has the child been ill during the past two weeks? ***(If no:Λ go to 3.09)***	***Y/N/D*** ~	3.04	How does caretaker consider the severity of that illness? ***(1=none, 2=mild, 3=moderate, 4=life threatening, 5=don't know)***	~
Self reported symptoms over the past 2 weeks *(Ranyisi mag tuo ni gin mage to kuom ndalo adi)* ***Only write in Dholuo - do not translate into English***					
3.05	*Symptom 1*	~~~	3.06	*Symptom 2*	~~~
3.07	*Symptom 3*	~~~	3.08	*Symptom 4*	~~~

Probe for symptoms in past 2 weeks		***Y/N/D***	**Probe for symptoms in past 2 weeks**		***Y/N/D***
3.09	del maore *fever*	~	3.10	kor mathung *difficulty getting air/problem with chest*	~
3.11	fudha *hot body/joint pains, yellow body*	~	3.12	lwedo marachar *white pale palms*	~
3.13	midhusi *serious type of malaria*	~	3.14	kokene rachar *white nails*	~
3.15	talarieya *seriously ill, fever, vomiting, convulsion*	~	3.16	remo matin *weak blood*	~
3.17	ndulume *untreated fudha, convulsions/ unconscious*	~	3.18	del monyosore *weak body*	~
3.19	ng'ok *vomiting*	~	3.20	del maraton'g *pale body*	~
3.21	dhok marach *loss of appetite*	~	3.22	wan'g marachar *pale eyes*	~
3.23	ok chiem/ dhoth *drinks poorly or unable to feed/nurse*	~	3.24	Lep marachar *tongue pallor*	~
3.25	dhero *weight loss*	~	3.26	odondwe *red eyes*	~
3.27	diep *diarrhea*	~	3.28	del maruodho *body rash*	~
3.29	diep mar remo *bloody diarrhea*	~	3.30	guonyruok *scratching/itching*	~
3.31	orianyanja *yellow/greenish diarrhoea*	~	3.32	it maremo/it machwer *ear pain/ infection*	~
3.33	diep mao ka pi *watery diarrhea*	~	3.34	athunga *runny nose*	~
3.35	ahonda *cough*	~	3.36	olo/oloch *very sleepy/listless*	~
3.37	thung *difficulty breathing*	~	3.38	odowa/lowo/chamloo *eating soil*	~
During the past <u>two months</u> (not weeks) ago did the child have any of the following illness (*Y/N/D*).					
3.39	diep *diarrhoea lasting 14 days or more*	~	3.40	talarieya *seriously ill, fever, vomit, convulsions*	~
3.41	midhusi *serious malaria*	~	3.42	ndulume *untreated fudha, convulsions, unconscious*	~

Part 4: Health Care Seeking and Treatment

If child has not been ill in the last 2 weeks, Λ go to 5.01					
4.01	Was the child taken anywhere to seek health care or get medicine in the last 2 weeks? (*If no; Λ go to 5.01*)				*Y/N/D* ~
4.02	Where was child taken first? (*1=private clinic 2=health centre/dispensary 3=hospital IPD [admitted] 4=hospital OPD 5=traditional healer/herbalist 6=bush doctor 7=shop/duka/chemist 8=CHW/nyamrerwas 9=market vendors/hawkers*)	~	4.03 How many days after the onset of symptoms? (*Day 0 is the first day of symptoms*)		~~
4.04	Was the child given any traditional medicine or any western/conventional medicine from this visit? (*1=traditional medicine 2=western medicine 3=no medicine*)	~ ~	4.05 If given western/conventional medicine, do you know the name? (*If yes;Λ go to 4.06, if no; Λ go to 4.10*) *Y/N*		~
4.06	*Medicine 1*	~~~	4.07	*Medicine 2*	~~~
4.08	*Medicine 3*	~~~	4.09	*Medicine 4*	~~~
4.10	Where was child taken next? (*1=private clinic 2=health centre/dispensary 3=hospital IPD [admitted] 4=hospital OPD 5=traditional healer/herbalist 6=bush doctor 7=shop/duka/chemist 8=CHW/nyamrerwas 9=market vendors/hawkers 10=nowhere*)	~~	4.11 How many days after the onset of symptoms? (*Day 0 is the first day of symptoms*)		~~
4.12	Was the child given any traditional medicine or any western/conventional medicine from this visit? (*1=traditional medicine 2=western medicine 3=no medicine*)	~	4.13 If given western/conventional medicine, do you know the name? (*If yes;Λ go to 4.14, if no; Λ go to 5.01*) *Y/N*		~
4.14	*Medicine 1*	~~~	4.15	*Medicine 2*	~~~
4.16	*Medicine 3*	~~~	4.17	*Medicine 4*	~~~

Part 5: Clinical Examination (Village Monitors)

5.01	Breathing frequency / 60 sec (*count for one full minute*)	~~~	5.02	Increased work of breathing? (*N=none, M=mild, S=severe*)	*N/M/S* ~
5.03	Palm pallor(*N=none, M=mild/moderate, S=severe*)	*N/M/S* ~	5.04	Eyelid pallor (*N=none, M=mild/moderate, S=severe*)	*N/M/S* ~
5.05	Tongue pallor (*N=none, M= mild/moderate, S=severe*)	*N/M/S* ~	5.06	Nail pallor *N=none, M= mild/moderate, S=severe*)	*N/M/S* ~

Part 6: Clinical measurements (Gestational Age Nyamrerwas)

6.01	Weighing scale number	~~~	*If <6 months use 10 kg scale*	6.02	Height board number	~~	*If < 23 months, take length lying down*
6.03	Weight I (*kg*)	~~.~~	*If difference >0.30 kg, weigh child again*	6.04	Length/height I (*cm*)	~~~.~	*If difference > 2 cm, measure child again*
6.05	Weight II (*kg*)	~~.~~		6.06	Length/height II (*cm*)	~~~.~	
6.07	MUAC I (*cm*)	~~.~	*If difference >1 cm, measure child again*	6.08	Axillary temperature I (*ºC) (under the arm-pit*)	~~.~~	*If temperature <36.0 take it again*
6.09	MUAC II (*cm*)	~~.~		6.10	Axillary temperature II (*ºC) (under the arm-pit*)	~~.~~	

Part 7: Clinical Examination (Clinical Officer)

7.01	Does the child seem ill? ***(1=no, 2=mild, 3=moderate, 4=life-threatening)***	~	7.02	Del maroudho/gwonyo ***(evidence of insect bite rash/scabies on lower arms or legs ?)***	***Y/N/D*** ~
7.03	Skin ***(1= normal, 2=thin)***	~	7.04	Visible severe wasting? ***(Wasted buttocks or bony thorax structure)***	***Y/N/D*** ~
7.05	Evidence of BCG scar	***Y/N/D*** ~	7.06	Hair colour ***(1=normal, 2=light)***	~
7.07	Hair texture ***(1=normal, 2=thin)***	~	7.08	Oedema in both feet? ***(1=none, 2=mild, 3=severe)***	~
7.09	Spleen (Hacket score) ***(0,1,2,3,4)***	~	7.10	Any handicaps? ***(1=no 2=congenital malformation 3=blind 4=deaf 5=crippled 6=mental handicap)***	~

Please indicate all diagnoses requiring treatment, other than malaria and anaemia.

7.11	*Other diagnosis 1*	~~~	7.12	*Other diagnosis 2*	~~~
7.13	*Other diagnosis 3*	~~~	7.14	*Other diagnosis 4*	~~~

Part 8: Hb and Medication given by Clinical Officer

8.01	Hb result field	~~.~	8.02	Fansidar	***Y/N*** ~
8.03	Amodiaquine	***Y/N*** ~			

Quality Control

	Date	Name
Supervisor	~~/~~/~~	
Data Entry Clerk	~~/~~/~~	

Appendix B. Laboratory procedures

DNA extraction: DNA was purified from red blood cell pellets using the QIAamp DNA Blood Mini Kit (Qiagen, CA, USA).

SNP genotyping of drug resistance markers: Real-time polymerase chain reaction (PCR) (Stratagene Mx3005P, CA, USA) was used to detect single nucleotide polymorphisms (SNPs) at *dhfr*-51, 59, 108 and 164, *dhps*-437 and 540, *pfcrt*-76 and *pfmdr1*-86 using published procedures (33-35). Briefly, standards and field samples were run in duplicate in 25 µL reactions containing TaqMan Universal Mastermix (Applied Biosystems, CA, USA), 2 µL of DNA (diluted 1:10), gene-specific forward and reverse primers, and SNP-specific TaqMan MGB probes (Applied Biosystems, CA, USA). Serial dilutions of both wild type and mutant parasite laboratory strain standards, depending on the SNP, and negative control templates were run on every plate as positive and negative controls.

Microsatellite (MS) typing of pfg377: The MS marker within the coding region of the *pfg377* gene was analyzed using a method described elsewhere (30). In short, amplifications were carried out using single reaction PCR with fluorescent labeled primers incorporated with FAM dye. PCR products were read on ABI (Applied Biosystems 3100) capillary sequencer and GeneMapper software was used to automate measurement of microsatellite base-pair length and quantify peak height. Multiple alleles were defined based on the presence of a minor allele peak height ≥30% of the predominant allele (30, 36).

Appendix C. Variable and outcome descriptions and coding

Explanatory variable	Description	Type/Coding
BEDNET	Reflects year 5 or baseline survey	Categorical (1=year and 0=baseline
ANTIFOLATE	Reflects SP and/or CTX (Septrin) within 2 weeks prior to survey	Categorical (1=SP and/or CTX use in past 2 weeks, 0=neither SP nor CTX use in past 2 weeks)
SP2W	Used SP within 2 weeks prior to survey, yes/no	Categorical (1=yes, 0=no)
CQ2W	Used CQ within 2 weeks prior to survey, yes/no	Categorical (1=yes, 0=no)
SEPTRIN	Used septrin (CTX) within 2 weeks prior to survey, yes/no	Categorical (1=yes, 0=no)
AGEMONTH	Age in months	Continuous
SEX	Male or Female	Categorical (1=male, 0=female)
FEVER	Temperature greater than 37.5C within 48 hrs prior to survey, yes/no	Categorical (1=yes, 0=no)
HB	Hemoglobin level based on HemoCue system, recorded as g/DL	Continuous
PMM3	Parasites per uL blood (parasite density)	Continuous
LNPMM3	Log transformed parasites per uL blood (parasite density) + 1	Continuous
GAMCYTES	Presence of gametocytes by microscopy, yes/no	Categorical (1=yes, 0=no)
PFGPREV	Diversity of *pfg377* microsatellite marker	Categorical (1=multiple infection, 0=single infection)
GDSHORE	GIS distance to shore (meters)	Continuous
NEWGDSHORE	GIS distance to shore (kilometers)	Continuous
GDCLINIC	GIS distance to nearest clinic (meters)	Continuous
NEWGDCLINIC	GIS distance to nearest clinic (kilometers)	Continuous
GDROADS	GIS distance to nearest compound (meters)	Continuous
LNROADS	Log transformed GIS distance to shore (meters)	Continuous
ELEVTION	Elevation of compound (meters)	Continuous

Genotype	Based on mutations in SNPs	Genotype Categories (# of categories)	Outcome Name	Outcome Categories	Type of Logistic Regression
DHFR	*dhfr*-51/59/108	Wild Type*, Single, Double, Triple (4)	*dhfr* triple mutant	1. WT, Single, Double (ref) 2. Triple	Binary
DHPS	*dhps*-437/540	Wild Type*, Single, Double (3)	*dhps* mutations	1. WT 2. Single 3. Double	Cumulative
Combined DHFR/DHPS	*dhfr*-51/59/108 + *dhps*-437/540	Wild Type*, Single, Double, Triple, Quadruple, Quintuple (6)	combined *dhfr/dhps* mutations	1. WT, Single, Double, Triple 2. Quadruple 3. Quintuple	Cumulative
PFCRT-76 or **PFMDR1-86**	*crt*-76; *mdr1*-86	Wild Type*, Mutant (2)	*pfcrt-76* mutant or *pfmdr1-86* mutant	1. WT (ref) 2. Mutant	Binary

SNPs, single nucleotide polymorphisms

*Wild Type, WT, refers to no mutations in any of the specified SNPs

Appendix D. Univariable analysis for all study variables for genes associated with SP resistance among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya

Variable	Drug resistance outcome		
	dhfr Triple Mutant[a] OR (95% CI)	*dhps* Mutations[b] OR (95% CI)	*dhfr/dhps* Mutations[c] OR (95% CI)
Survey Year[d]	1.62 (1.09, 2.41)*	9.21 (6.19, 13.71)*	7.48 (5.17, 10.83)*
Antifolate use[e,f]	1.09 (0.57, 2.09)	3.17 (1.63, 6.15)*	2.16 (1.24, 3.78)*
SP use[f,g]	1.01 (0.38, 2.65)	5.62 (1.63, 19.42)*	2.98 (1.21, 7.31)*
CTX use[f]	1.07 (0.48, 2.36)	2.33 (1.08, 5.03)*	1.66 (0.85, 3.25)
Age (months)	1.02 (1.00, 1.03)*	1.03 (1.02, 1.04)*	1.02 (1.01, 1.03)*
Sex	0.87 (0.59, 1.28)	1.04 (0.74, 1.46)	0.87 (0.63, 1.21)
Fever[h]	1.40 (0.67, 2.92)	0.84 (0.47, 1.50)	0.92 (0.52, 1.62)
GIS distance to clinic (km)	1.19 (0.98, 1.45)	1.13 (0.95, 1.33)	1.15 (0.98, 1.35)
GIS distance to nearest compound (m)[i]	1.08 (0.88, 1.32)	0.98 (0.82, 1.17)	0.98 (0.82, 1.16)
GIS distance to lake shore (km)	0.94 (0.87, 1.01)	0.98 (0.92, 1.05)	0.96 (0.91, 1.03)
GIS elevation (m)	1.00 (0.99, 1.00)	1.00 (1.00, 1.00)	1.00 (0.99, 1.00)
Hemoglobin (g/dL)	1.01 (0.93, 1.11)	1.23 (1.16, 1.36)*	1.22 (1.13, 1.31)*
Parasite density (uL)[i]	1.02 (0.91, 1.14)	0.93 (0.84, 1.03)	0.94 (0.86, 1.04)
Pfg377 microsatellite diversity[j]	2.34 (1.54, 3.55)*	1.27 (0.89, 1.79)	1.70 (1.22, 2.38)*
Gametocytes	1.23 (0.74, 2.03)	1.07 (0.70, 1.64)	1.22 (0.81, 1.83)

Abbreviations: OR, odds ratio; CI, confidence interval; MS, microsatellite; SP, sulfadoxine-pyrimethamine; CTX, cotrimoxazole.

[a] *dhfr* Triple Mutant was analyzed by grouping wild type, single and double genotypes as the reference category.

[b] *dhps* mutations were analyzed as 3 genotype categories: (1) wild type, (2) single, and (3) double.

[c] *dhfr* and *dhps* combined mutations were analyzed as 3 genotype categories: (1) wild type, single, double, and triple, (2) quadruple, and (3) quintuple.

[d] Year 5 survey compared to baseline.

[e] Model containing SP and/or CTX use as predictor.

[f] Within two weeks prior to survey.

[g] Model containing SP use only as predictor.
[h] Body temperature greater than or equal to 37.5°C at or 48 hours before survey.
[i] Log transformed prior to statistical analysis.
[j] Infection with more than one allele based on microsatellite marker in coding region of *pfg377*.
* Statistically significant, $p<0.05$.

Appendix E. Univariable analysis for all study variables for genes associated with CQ resistance among *P. falciparum* smear-positive children under the age of five during the ITN trial in western Kenya

Variable	Drug resistance outcome	
	pfcrt-76 Mutant[a] OR (95% CI)	*pfmdr1-86* Mutant[a] OR (95% CI)
Survey Year[b]	1.02 (0.64, 1.60)	0.92 (0.61, 1.38)
CQ use[c]	0.98 (0.51, 1.87)	1.46 (0.78, 2.72)
Age (months)	1.00 (0.98, 1.01)	1.01 (1.00, 1.02)
Sex	1.19 (0.75, 1.88)	1.15 (0.77, 1.72)
Fever[d]	0.87 (0.40, 1.89)	1.04 (0.51, 2.13)
GIS distance to clinic	1.05 (0.83, 1.31)	0.97 (0.80, 1.19)
GIS distance to nearest compound[e]	1.19 (0.95, 1.50)	1.04 (0.85, 1.29)
GIS distance to lake shore	1.00 (0.92, 1.10)	0.95 (0.88, 1.03)
GIS elevation	1.00 (1.00, 1.00)	1.00 (0.99, 1.00)
Hemoglobin (g/dL)	1.00 (0.91, 1.11)	1.00 (0.92, 1.10)
Parasite density (uL) [e]	0.95 (0.83, 1.09)	1.06 (0.94, 1.19)
Pfg377 microsatellite diversity[f]	0.60 (0.38, 0.96)*	1.38 (0.91, 2.09)
Gametocytes	1.46 (0.78, 2.70)	0.97 (0.58, 1.59)

Abbreviations: OR, odds ratio; CI, confidence interval; MS, microsatellite; CQ, chloroquine.

[a] *pfcrt*-76 and *pfmdr1*-86 mutants were analyzed using wild type as the reference group.

[b] Year 5 survey compared to baseline.

[c] Within two weeks prior to survey.

[d] Body temperature greater than or equal to 37.5°C at or 48 hours before survey.

[e] Log transformed prior to statistical analysis.

[f] Infection with more than one allele based on microsatellite marker in coding region of *pfg377*.

* Statistically significant, $p<0.05$.

Appendix F. Assessment of Interaction

E_1 = SP use (yes/no)
E_2 = Antifolate use (yes/no)
E_3 = CQ use (yes/no)
E_4 = Survey year (2001 or 1996)
V_1 =*pfg377* MS diversity (>1 or ≤1 allele)
V_2 = GIS distance to lake shore (km)
V_3 =Sex (male/female)
V_4 = Age (months)
V_5 =Hemoglobin (g/dL)
V_6 = Parasite density (uL)
V_7 = Gametocytes present (yes/no)
V_8 = Fever in last 48 hours (yes/no)
V_9 = GIS distance to nearest clinic (km)

SP Model 1 (Full): E_1, E_4, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, E_1*E_4, E_1*V_1, E_1*V_2, E_1*V_3, E_1*V_4, E_1*V_5, E_1*V_6, E_1*V_7, E_1*V_8, E_1*V_9, E_4*V_1, E_4*V_2, E_4*V_3, E_4*V_4, E_4*V_5, E_4*V_6, E_4*V_7, E_4*V_8, E_4*V_9

SP Model 2 (Reduced): E_1, E_4, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9

Antifolate Model 1 (Full): E_2, E_4, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, E_2*E_4, E_2*V_1, E_2*V_2, E_2*V_3, E_2*V_4, E_2*V_5, E_2*V_6, E_2*V_7, E_2*V_8, E_2*V_9, E_4*V_1, E_4*V_2, E_4*V_3, E_4*V_4, E_4*V_5, E_4*V_6, E_4*V_7, E_4*V_8, E_4*V_9

Antifolate Model 2 (Reduced): E_2, E_4, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9

CQ Model 1 (Full): E_3, E_4, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, E_3*E_4, E_3*V_1, E_3*V_2, E_3*V_3, E_3*V_4, E_3*V_5, E_3*V_6, E_3*V_7, E_3*V_8, E_3*V_9, E_4*V_1, E_4*V_2, E_4*V_3, E_4*V_4, E_4*V_5, E_4*V_6, E_4*V_7, E_4*V_8, E_4*V_9

CQ Model 2 (Reduced): E_3, E_4, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9

Model	Outcome	-2LogL	Likelihood ratio test statistic (degrees of freedom)*	p-value
Full (SP Model 1)	*dhfr* Triple Mutant	480.310	21.027 (19)	0.34
Reduced (SP Model 2)	*dhfr* Triple Mutant	501.337		
Full (SP Model 1)	*dhps* Mutations	741.368	17.013 (19)	0.59
Reduced (SP Model 2)	*dhps* Mutations	758.381		
Full (SP Model 1)	*dhfr/dhps* Mutations	826.109	13.925 (19)	0.79
Reduced (SP Model 2)	*dhfr/dhps* Mutations	840.034		
Full (Antifolate Model 1)	*dhfr* Triple Mutant	488.253	21.364 (19)	0.32
Reduced (Antifolate Model 2)	*dhfr* Triple Mutant	509.617		
Full (Antifolate Model 1)	*dhps* Mutations	753.421	17.620 (19)	0.55
Reduced (Antifolate Model 2)	*dhps* Mutations	771.041		
Full (Antifolate Model 1)	*dhfr/dhps* Mutations	843.651	11.827 (19)	0.89
Reduced (Antifolate Model 2)	*dhfr/dhps* Mutations	855.478		
Full (CQ Model 1)	*pfcrt*-76 Mutant	412.854	17.685 (19)	0.54
Reduced (CQ Model 2)	*pfcrt*-76 Mutant	430.539		
Full (CQ Model 1)	*pfmdr1*-86 Mutant	478.195	17.58	0.55
Reduced (CQ Model 2)	*pfmdr1*-86 Mutant	495.775		

*Test statistic is distributed chi-square under null hypothesis

Appendix G. Collinearity Information Matrices

CHLOROQUINE RESISTANCE

(Same explanatory variables for outcomes: pfmdr186 Mutant and pfcrt76 Mutant)

```
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model mdr186=newgdshore elevtion lnroads lnpmm3 agemonth sex hb
      bednet cq2w pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES, AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6	VDP7
EIGENVAL	0.000	0.0104	0.0280	0.0425	0.0876	0.11359	0.16414
CONDINDX	309.763	30.1584	18.4287	14.9506	10.4159	9.14519	7.60777
	.	.	.	.	.	.	.
Intercept	0.993	0.0070	0.0001	0.0000	0.0001	0.00002	0.00000
newgdshore	0.843	0.0413	0.0005	0.0012	0.0799	0.03228	0.00009
elevtion	0.995	0.0048	0.0000	0.0000	0.0000	0.00004	0.00000
lnroads	0.035	0.5691	0.3108	0.0436	0.0375	0.00173	0.00027
lnpmm3	0.004	0.2149	0.1702	0.5338	0.0650	0.00245	0.00102
agemonth	0.001	0.0037	0.0732	0.1837	0.0468	0.13851	0.50903
sex	0.000	0.0095	0.0022	0.0016	0.0076	0.00061	0.00745
hb	0.000	0.1451	0.5883	0.2566	0.0046	0.00003	0.00378
bednet	0.035	0.0060	0.0116	0.0477	0.0607	0.11775	0.28513
cq2w	0.001	0.0006	0.0106	0.0009	0.0155	0.00813	0.00114
pfgprev	0.000	0.0107	0.0011	0.0117	0.0052	0.06257	0.00065
gamcytes	0.000	0.0004	0.0063	0.0077	0.0011	0.00506	0.01030
fever	0.000	0.0073	0.0140	0.0209	0.0303	0.00695	0.00976
newgdclinic	0.000	0.0155	0.0169	0.0000	0.1322	0.60626	0.20592

VARIABLE	VDP8	VDP9	VDP10	VDP11	VDP12	VDP13	VDP14
EIGENVAL	0.41316	0.49526	0.53097	0.76524	0.89191	0.95720	9.49993
CONDINDX	4.79512	4.37968	4.22984	3.52339	3.26362	3.15035	1.00000
	.	.	.	.	.	.	.
Intercept	0.00002	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
newgdshore	0.00113	0.00002	0.00000	0.00001	0.00002	0.00001	0.00013
elevtion	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
lnroads	0.00168	0.00024	0.00000	0.00012	0.00005	0.00002	0.00025
lnpmm3	0.00697	0.00049	0.00040	0.00001	0.00000	0.00018	0.00045
agemonth	0.01565	0.00001	0.02113	0.00228	0.00013	0.00333	0.00132
sex	0.38276	0.13563	0.42703	0.01729	0.00337	0.00174	0.00295
hb	0.00007	0.00018	0.00066	0.00019	0.00001	0.00009	0.00036
bednet	0.24503	0.03206	0.14173	0.00034	0.00161	0.01304	0.00200
cq2w	0.11129	0.00523	0.00345	0.01414	0.37311	0.45363	0.00110
pfgprev	0.00191	0.74656	0.14965	0.00087	0.00578	0.00031	0.00288
gamcytes	0.00445	0.00260	0.01424	0.90465	0.00185	0.03918	0.00213
fever	0.00682	0.00985	0.02494	0.02784	0.53234	0.30782	0.00105
newgdclinic	0.01248	0.00555	0.00255	0.00036	0.00013	0.00044	0.00131

```
*drop elevtion;
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model mdr186=newgdshore lnroads lnpmm3 agemonth sex hb bednet
      cq2w pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6	VDP7
EIGENVAL	0.0069	0.0278	0.0422	0.08680	0.11076	0.16414	0.39768
CONDINDX	35.0955	17.5161	14.2098	9.91271	8.77549	7.20873	4.63121
	.	.	.	.	.	.	.
Intercept	0.9886	0.0027	0.0020	0.00435	0.00109	0.00000	0.00096
newgdshore	0.1396	0.0083	0.0055	0.51743	0.31563	0.00064	0.01125
lnroads	0.5429	0.3570	0.0511	0.04337	0.00214	0.00028	0.00247
lnpmm3	0.2446	0.1609	0.5011	0.07778	0.00394	0.00099	0.00920
agemonth	0.0020	0.0665	0.1945	0.05613	0.12826	0.51082	0.01286
sex	0.0115	0.0023	0.0017	0.00832	0.00067	0.00750	0.33583
hb	0.1549	0.5533	0.2803	0.00571	0.00001	0.00382	0.00028
bednet	0.0047	0.0109	0.0500	0.06964	0.10176	0.29770	0.27998
cq2w	0.0018	0.0110	0.0010	0.01717	0.00758	0.00116	0.10806
pfgprev	0.0087	0.0007	0.0099	0.00868	0.06562	0.00066	0.00042
gamcytes	0.0004	0.0062	0.0072	0.00060	0.00633	0.01028	0.00582
fever	0.0073	0.0131	0.0188	0.03392	0.00653	0.00975	0.00666
newgdclinic	0.0221	0.0162	0.0000	0.16374	0.56296	0.20408	0.02028

VARIABLE	VDP8	VDP9	VDP10	VDP11	VDP12	VDP13
EIGENVAL	0.49202	0.53097	0.76313	0.89060	0.95737	8.52956
CONDINDX	4.16364	4.00801	3.34320	3.09473	2.98486	1.00000
	.	.	.	.	.	.
Intercept	0.00013	0.00000	0.00004	0.00001	0.00000	0.00012
newgdshore	0.00012	0.00002	0.00010	0.00015	0.00008	0.00119
lnroads	0.00022	0.00000	0.00014	0.00005	0.00002	0.00031
lnpmm3	0.00036	0.00040	0.00001	0.00000	0.00018	0.00055
agemonth	0.00012	0.02105	0.00267	0.00017	0.00329	0.00165
sex	0.17815	0.42654	0.01885	0.00316	0.00166	0.00369
hb	0.00023	0.00066	0.00023	0.00001	0.00009	0.00045
bednet	0.02132	0.14597	0.00068	0.00139	0.01332	0.00262
cq2w	0.00173	0.00330	0.01063	0.38370	0.45154	0.00136
pfgprev	0.74237	0.15244	0.00141	0.00514	0.00035	0.00361
gamcytes	0.00238	0.01429	0.90214	0.00306	0.03865	0.00269
fever	0.01017	0.02506	0.03073	0.52489	0.31174	0.00133
newgdclinic	0.00534	0.00259	0.00043	0.00016	0.00043	0.00162

```
*drop lnroads;
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model mdr186=newgdshore lnpmm3 agemonth sex hb bednet cq2w
      pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6
EIGENVAL	0.0119	0.0397	0.0750	0.11009	0.16384	0.38042
CONDINDX	25.2383	13.8260	10.0616	8.30504	6.80765	4.46759
	.	.	.	.	.	.
Intercept	0.9702	0.0088	0.0162	0.00174	0.00001	0.00241
newgdshore	0.0906	0.0303	0.4480	0.41114	0.00020	0.01761
lnpmm3	0.4995	0.2650	0.2181	0.00310	0.00082	0.01197
agemonth	0.0053	0.2752	0.0492	0.10708	0.52311	0.00976
sex	0.0062	0.0056	0.0133	0.00049	0.00794	0.30026
hb	0.4284	0.5498	0.0152	0.00002	0.00410	0.00069
bednet	0.0099	0.0492	0.0723	0.08077	0.31205	0.29657
cq2w	0.0191	0.0002	0.0131	0.00521	0.00098	0.09464
pfgprev	0.0048	0.0051	0.0197	0.06282	0.00072	0.00421
gamcytes	0.0061	0.0026	0.0005	0.00763	0.00959	0.00982
fever	0.0168	0.0052	0.0537	0.00459	0.01019	0.00629
newgdclinic	0.0104	0.0112	0.2376	0.50664	0.19080	0.03239

VARIABLE	VDP7	VDP8	VDP9	VDP10	VDP11	VDP12
EIGENVAL	0.49009	0.53067	0.75970	0.88943	0.95609	7.59303
CONDINDX	3.93614	3.78263	3.16146	2.92181	2.81811	1.00000
	.	.	.	.	.	.
Intercept	0.00021	0.00000	0.00008	0.00003	0.00000	0.00027
newgdshore	0.00010	0.00002	0.00016	0.00019	0.00007	0.00155
lnpmm3	0.00027	0.00040	0.00002	0.00000	0.00019	0.00069
agemonth	0.00047	0.02125	0.00317	0.00019	0.00318	0.00209
sex	0.20898	0.42675	0.02131	0.00301	0.00144	0.00471
hb	0.00027	0.00066	0.00027	0.00002	0.00008	0.00056
bednet	0.01334	0.14761	0.00121	0.00118	0.01255	0.00337
cq2w	0.00019	0.00351	0.00566	0.41117	0.44457	0.00170
pfgprev	0.73970	0.15099	0.00240	0.00441	0.00057	0.00464
gamcytes	0.00173	0.01490	0.90196	0.00522	0.03644	0.00350
fever	0.01081	0.02515	0.03437	0.50257	0.32861	0.00172
newgdclinic	0.00524	0.00260	0.00053	0.00018	0.00041	0.00205

SULFADOXINE-PYRAMETHAMINE RESISTANCE

(Same explanatory variables for outcomes: dhfr triple mutant, dhps mutations, and combined dhfr/dhps mutations)

ANTIFOLATE USE as drug use predictor:

```
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model dhfroutcome=newgdshore elevtion lnroads lnpmm3 agemonth sex
hb bednet antifolate pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES, AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6	VDP7
EIGENVAL	0.000	0.0101	0.0283	0.0419	0.0895	0.11796	0.16796
CONDINDX	310.286	30.4945	18.1599	14.9418	10.2221	8.90240	7.46070
	.	.	.	.	.	.	.
Intercept	0.993	0.0072	0.0001	0.0000	0.0001	0.00003	0.00000
newgdshore	0.848	0.0423	0.0000	0.0005	0.0849	0.02354	0.00026
elevtion	0.995	0.0049	0.0000	0.0000	0.0000	0.00004	0.00000
lnroads	0.033	0.5722	0.3331	0.0213	0.0341	0.00382	0.00057
lnpmm3	0.002	0.2123	0.2396	0.4861	0.0495	0.00340	0.00028
agemonth	0.002	0.0066	0.0586	0.1661	0.0658	0.18581	0.47926
sex	0.000	0.0082	0.0019	0.0009	0.0016	0.00029	0.00594
hb	0.000	0.1606	0.4887	0.3378	0.0068	0.00011	0.00411
bednet	0.032	0.0115	0.0063	0.0569	0.0685	0.13863	0.24929
antifolate	0.000	0.0200	0.0000	0.0003	0.0032	0.00000	0.00001
pfgprev	0.000	0.0069	0.0037	0.0084	0.0002	0.08368	0.00152
gamcytes	0.001	0.0004	0.0128	0.0024	0.0033	0.01180	0.01129
fever	0.000	0.0091	0.0135	0.0141	0.0139	0.01167	0.00086
newgdclinic	0.001	0.0187	0.0253	0.0001	0.0703	0.60446	0.25903

VARIABLE	VDP8	VDP9	VDP10	VDP11	VDP12	VDP13	VDP14
EIGENVAL	0.43380	0.53148	0.56129	0.76600	0.90086	1.00203	9.34878
CONDINDX	4.64228	4.19406	4.08115	3.49352	3.22143	3.05448	1.00000
	.	.	.	.	.	.	.
Intercept	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000
newgdshore	0.00060	0.00011	0.00003	0.00001	0.00005	0.00001	0.00013
elevtion	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000
lnroads	0.00113	0.00067	0.00011	0.00005	0.00005	0.00006	0.00025
lnpmm3	0.00302	0.00140	0.00131	0.00000	0.00000	0.00023	0.00044
agemonth	0.00343	0.00037	0.02690	0.00233	0.00097	0.00036	0.00138
sex	0.80694	0.00886	0.15378	0.00549	0.00133	0.00122	0.00308
hb	0.00015	0.00028	0.00048	0.00009	0.00009	0.00000	0.00038
bednet	0.06874	0.06395	0.29525	0.00126	0.00019	0.00579	0.00209
antifolate	0.01889	0.07719	0.07844	0.18692	0.26643	0.34733	0.00128
pfgprev	0.02106	0.71567	0.09600	0.01483	0.01541	0.02985	0.00282
gamcytes	0.00896	0.00363	0.00112	0.78498	0.02640	0.12972	0.00192
fever	0.00547	0.00004	0.02062	0.00885	0.61484	0.28600	0.00084
newgdclinic	0.00599	0.01285	0.00019	0.00046	0.00047	0.00003	0.00142

```
*drop elevtion;
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model dhfroutcome=newgdshore lnroads lnpmm3 agemonth sex hb
bednet antifolate pfgprev
      gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6	VDP7
EIGENVAL	0.0067	0.0282	0.0417	0.08916	0.11444	0.16786	0.42382
CONDINDX	35.4495	17.2469	14.1791	9.69518	8.55758	7.06584	4.44678
	.	.	.	.	.	.	.
Intercept	0.9899	0.0028	0.0012	0.00348	0.00146	0.00001	0.00059
newgdshore	0.1510	0.0019	0.0015	0.58847	0.24582	0.00174	0.00658
lnroads	0.5505	0.3768	0.0248	0.03932	0.00483	0.00059	0.00180
lnpmm3	0.2395	0.2301	0.4587	0.05761	0.00566	0.00026	0.00463
agemonth	0.0044	0.0522	0.1757	0.07610	0.17426	0.48317	0.00322
sex	0.0098	0.0021	0.0009	0.00156	0.00009	0.00613	0.75083
hb	0.1681	0.4561	0.3620	0.00787	0.00000	0.00416	0.00036
bednet	0.0097	0.0055	0.0590	0.07627	0.12274	0.26220	0.10193
antifolate	0.0199	0.0000	0.0004	0.00298	0.00003	0.00002	0.01391
pfgprev	0.0055	0.0031	0.0070	0.00119	0.08841	0.00139	0.00640
gamcytes	0.0003	0.0119	0.0018	0.00251	0.01315	0.01152	0.00939
fever	0.0086	0.0127	0.0129	0.01626	0.01235	0.00088	0.00683
newgdclinic	0.0252	0.0246	0.0000	0.09286	0.57409	0.25483	0.01160

VARIABLE	VDP8	VDP9	VDP10	VDP11	VDP12	VDP13
EIGENVAL	0.52214	0.55936	0.76622	0.89990	1.00000	8.38058
CONDINDX	4.00632	3.87073	3.30719	3.05169	2.89493	1.00000
	.	.	.	.	.	.
Intercept	0.00032	0.00004	0.00002	0.00002	0.00001	0.00012
newgdshore	0.00102	0.00016	0.00006	0.00040	0.00011	0.00123
lnroads	0.00077	0.00007	0.00006	0.00006	0.00006	0.00031
lnpmm3	0.00172	0.00105	0.00000	0.00000	0.00024	0.00054
agemonth	0.00059	0.02453	0.00264	0.00118	0.00031	0.00174
sex	0.05291	0.16312	0.00582	0.00166	0.00132	0.00387
hb	0.00030	0.00050	0.00010	0.00011	0.00000	0.00048
bednet	0.08732	0.26490	0.00199	0.00043	0.00532	0.00275
antifolate	0.06769	0.09296	0.18214	0.27385	0.34444	0.00163
pfgprev	0.65941	0.16242	0.01302	0.01719	0.03148	0.00355
gamcytes	0.00343	0.00033	0.79201	0.02709	0.12401	0.00244
fever	0.00002	0.02126	0.00768	0.60207	0.29732	0.00107
newgdclinic	0.01346	0.00055	0.00046	0.00053	0.00003	0.00176

```
*drop lnroads;
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model dhfroutcome=newgdshore lnpmm3 agemonth sex hb bednet
      antifolate pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6
EIGENVAL	0.0117	0.0405	0.07680	0.11271	0.16720	0.40814
CONDINDX	25.2599	13.5603	9.84930	8.13044	6.67521	4.27251
	.	.	.	.	.	.
Intercept	0.9737	0.0049	0.01622	0.00235	0.00004	0.00175
newgdshore	0.1092	0.0098	0.47707	0.38788	0.00083	0.01156
lnpmm3	0.5240	0.2850	0.17586	0.00425	0.00017	0.00714
agemonth	0.0081	0.2320	0.09077	0.12625	0.50999	0.00260
sex	0.0040	0.0024	0.00510	0.00002	0.00716	0.67312
hb	0.4195	0.5501	0.02326	0.00000	0.00456	0.00080
bednet	0.0137	0.0552	0.08665	0.08116	0.28648	0.14052
antifolate	0.0100	0.0001	0.00427	0.00014	0.00001	0.00859
pfgprev	0.0051	0.0034	0.01152	0.08623	0.00098	0.00021
gamcytes	0.0082	0.0002	0.00004	0.01621	0.01091	0.01256
fever	0.0166	0.0053	0.03153	0.00903	0.00101	0.00798
newgdclinic	0.0074	0.0061	0.20673	0.51014	0.22840	0.02304

VARIABLE	VDP7	VDP8	VDP9	VDP10	VDP11	VDP12
EIGENVAL	0.51313	0.55927	0.76559	0.89861	0.99604	7.45032
CONDINDX	3.81043	3.64987	3.11952	2.87940	2.73494	1.00000
	.	.	.	.	.	.
Intercept	0.00060	0.00006	0.00003	0.00003	0.00002	0.00028
newgdshore	0.00121	0.00014	0.00003	0.00050	0.00015	0.00161
lnpmm3	0.00174	0.00091	0.00000	0.00000	0.00026	0.00067
agemonth	0.00035	0.02309	0.00313	0.00133	0.00023	0.00221
sex	0.12152	0.17102	0.00696	0.00216	0.00157	0.00495
hb	0.00035	0.00051	0.00013	0.00013	0.00001	0.00060
bednet	0.08361	0.24104	0.00292	0.00074	0.00445	0.00356
antifolate	0.06986	0.09926	0.17104	0.29503	0.33952	0.00215
pfgprev	0.62329	0.19960	0.01070	0.02157	0.03291	0.00455
gamcytes	0.00302	0.00006	0.80469	0.02531	0.11552	0.00321
fever	0.00007	0.02015	0.00816	0.57799	0.32080	0.00137
newgdclinic	0.01396	0.00080	0.00054	0.00058	0.00001	0.00223

SULFADOXINE-PYRAMETHAMINE RESISTANCE

(Same explanatory variables for outcomes: dhfr triple mutant, dhps mutations, and combined dhfr/dhps mutations)

SP USE as drug use predictor:

```
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model dhfroutcome=newgdshore elevtion lnroads lnpmm3 agemonth sex
      hb bednet sp2w pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6	VDP7
EIGENVAL	0.000	0.0100	0.0284	0.0420	0.0896	0.11940	0.16811
CONDINDX	308.538	30.5042	18.0817	14.8696	10.1836	8.82269	7.43545
	.	.	.	.	.	.	.
Intercept	0.993	0.0072	0.0001	0.0000	0.0001	0.00003	0.00000
newgdshore	0.848	0.0417	0.0000	0.0005	0.0864	0.02256	0.00019
elevtion	0.995	0.0049	0.0000	0.0000	0.0000	0.00004	0.00000
lnroads	0.036	0.5834	0.3187	0.0221	0.0333	0.00392	0.00037
lnpmm3	0.003	0.2064	0.2364	0.4944	0.0488	0.00335	0.00049
agemonth	0.001	0.0065	0.0639	0.1665	0.0542	0.17780	0.49572
sex	0.000	0.0070	0.0017	0.0020	0.0030	0.00013	0.00828
hb	0.000	0.1612	0.4999	0.3268	0.0063	0.00007	0.00405
bednet	0.033	0.0115	0.0057	0.0622	0.0629	0.14190	0.23532
sp2w	0.000	0.0251	0.0005	0.0051	0.0003	0.00207	0.00007
pfgprev	0.000	0.0082	0.0056	0.0090	0.0000	0.08391	0.00090
gamcytes	0.001	0.0000	0.0134	0.0036	0.0037	0.00988	0.01157
fever	0.000	0.0111	0.0138	0.0134	0.0138	0.01056	0.00116
newgdclinic	0.000	0.0167	0.0251	0.0002	0.0743	0.61037	0.25239

VARIABLE	VDP8	VDP9	VDP10	VDP11	VDP12	VDP13	VDP14
EIGENVAL	0.43432	0.53178	0.56129	0.79799	0.91593	1.00693	9.29409
CONDINDX	4.62594	4.18059	4.06923	3.41276	3.18546	3.03812	1.00000
	.	.	.	.	.	.	.
Intercept	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000
newgdshore	0.00070	0.00008	0.00000	0.00000	0.00000	0.00002	0.00013
elevtion	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000
lnroads	0.00116	0.00065	0.00004	0.00015	0.00004	0.00003	0.00025
lnpmm3	0.00353	0.00226	0.00014	0.00004	0.00030	0.00001	0.00045
agemonth	0.00387	0.00895	0.01649	0.00129	0.00219	0.00000	0.00141
sex	0.77993	0.12013	0.06464	0.00892	0.00062	0.00021	0.00312
hb	0.00013	0.00000	0.00075	0.00011	0.00001	0.00002	0.00039
bednet	0.08270	0.24179	0.10323	0.00036	0.01645	0.00128	0.00211
sp2w	0.01654	0.01271	0.09132	0.00390	0.29095	0.55078	0.00064
pfgprev	0.00949	0.32432	0.52964	0.00849	0.01351	0.00401	0.00292
gamcytes	0.00290	0.00176	0.02187	0.86120	0.02318	0.04424	0.00195
fever	0.00852	0.00709	0.02614	0.08117	0.55046	0.26189	0.00087
newgdclinic	0.00705	0.00558	0.00587	0.00034	0.00010	0.00035	0.00144

```
*drop elevtion;
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model dhfroutcome=newgdshore lnroads lnpmm3 agemonth sex hb
      bednet sp2w pfgprev gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6	VDP7
EIGENVAL	0.0066	0.0282	0.0418	0.08934	0.11572	0.16801	0.42345
CONDINDX	35.4520	17.1695	14.1083	9.65375	8.48214	7.03948	4.43414
	.	.	.	.	.	.	.
Intercept	0.9895	0.0029	0.0014	0.00349	0.00152	0.00001	0.00064
newgdshore	0.1476	0.0024	0.0018	0.60187	0.23569	0.00119	0.00734
lnroads	0.5636	0.3635	0.0260	0.03845	0.00502	0.00037	0.00184
lnpmm3	0.2357	0.2276	0.4651	0.05681	0.00582	0.00048	0.00528
agemonth	0.0043	0.0566	0.1767	0.06262	0.16912	0.49803	0.00356
sex	0.0086	0.0018	0.0020	0.00289	0.00002	0.00849	0.71104
hb	0.1691	0.4655	0.3522	0.00741	0.00000	0.00406	0.00033
bednet	0.0095	0.0049	0.0643	0.06979	0.12750	0.24668	0.12074
sp2w	0.0240	0.0008	0.0054	0.00038	0.00171	0.00007	0.01621
pfgprev	0.0067	0.0048	0.0074	0.00035	0.08942	0.00083	0.00144
gamcytes	0.0000	0.0126	0.0028	0.00294	0.01109	0.01179	0.00392
fever	0.0106	0.0129	0.0121	0.01620	0.01146	0.00119	0.01040
newgdclinic	0.0224	0.0243	0.0000	0.09649	0.57928	0.25003	0.01280

VARIABLE	VDP8	VDP9	VDP10	VDP11	VDP12	VDP13
EIGENVAL	0.52284	0.56036	0.79652	0.91479	1.00657	8.32572
CONDINDX	3.99050	3.85458	3.23305	3.01683	2.87601	1.00000
	.	.	.	.	.	.
Intercept	0.00031	0.00000	0.00004	0.00000	0.00001	0.00012
newgdshore	0.00061	0.00003	0.00006	0.00001	0.00018	0.00125
lnroads	0.00070	0.00003	0.00016	0.00004	0.00004	0.00031
lnpmm3	0.00219	0.00019	0.00005	0.00030	0.00002	0.00055
agemonth	0.00584	0.01764	0.00156	0.00218	0.00000	0.00177
sex	0.18292	0.06738	0.01003	0.00062	0.00032	0.00392
hb	0.00003	0.00076	0.00014	0.00001	0.00003	0.00048
bednet	0.21535	0.12119	0.00011	0.01620	0.00101	0.00279
sp2w	0.00625	0.09637	0.00710	0.29124	0.54958	0.00082
pfgprev	0.35632	0.49888	0.01119	0.01453	0.00445	0.00369
gamcytes	0.00061	0.02165	0.86560	0.01699	0.04748	0.00248
fever	0.00487	0.02880	0.07103	0.55916	0.26022	0.00111
newgdclinic	0.00666	0.00538	0.00037	0.00011	0.00039	0.00179

```
*drop lnroads;
%include 'C:\Users\Monica Shah\Documents\MPH Thesis\CDC folder
3.21.2011\collingenmodv9c.sas';
proc logistic data=genetic_clinic_only_v3 covout outtest=info;
      model dhfroutcome=newgdshore lnpmm3 agemonth sex hb bednet sp2w
pfgprev
      gamcytes fever newgdclinic / covb;
run;
%collin(covdsn=info);
run;
```

COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING
THE INFORMATION MATRIX: EIGENVALUES, CONDITION INDEXES,
AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)

VARIABLE	VDP1	VDP2	VDP3	VDP4	VDP5	VDP6
EIGENVAL	0.0118	0.0406	0.07699	0.11382	0.16754	0.40709
CONDINDX	25.0001	13.5042	9.80181	8.06121	6.64427	4.26249
	.	.	.	.	.	.
Intercept	0.9724	0.0056	0.01655	0.00251	0.00002	0.00190
newgdshore	0.1039	0.0113	0.49026	0.37887	0.00061	0.01236
lnpmm3	0.5230	0.2846	0.17622	0.00456	0.00038	0.00792
agemonth	0.0090	0.2362	0.07832	0.12513	0.52005	0.00278
sex	0.0042	0.0042	0.00671	0.00001	0.00930	0.62357
hb	0.4249	0.5464	0.02195	0.00000	0.00435	0.00075
bednet	0.0123	0.0593	0.08448	0.08622	0.26634	0.16355
sp2w	0.0054	0.0036	0.00016	0.00235	0.00010	0.01376
pfgprev	0.0078	0.0032	0.00928	0.08919	0.00056	0.00048
gamcytes	0.0071	0.0005	0.00003	0.01443	0.01129	0.00749
fever	0.0186	0.0045	0.03107	0.00826	0.00128	0.01173
newgdclinic	0.0063	0.0061	0.20997	0.51031	0.22734	0.02405

VARIABLE	VDP7	VDP8	VDP9	VDP10	VDP11	VDP12
EIGENVAL	0.51542	0.56006	0.79183	0.91186	1.00658	7.39641
CONDINDX	3.78819	3.63406	3.05630	2.84804	2.71072	1.00000
	.	.	.	.	.	.
Intercept	0.00057	0.00000	0.00008	0.00001	0.00002	0.00029
newgdshore	0.00066	0.00003	0.00012	0.00001	0.00023	0.00165
lnpmm3	0.00196	0.00024	0.00007	0.00032	0.00002	0.00069
agemonth	0.00313	0.01912	0.00192	0.00206	0.00001	0.00225
sex	0.25760	0.07572	0.01264	0.00067	0.00042	0.00502
hb	0.00008	0.00078	0.00018	0.00001	0.00004	0.00061
bednet	0.16998	0.13846	0.00000	0.01505	0.00062	0.00363
sp2w	0.00114	0.10443	0.01347	0.30666	0.54775	0.00112
pfgprev	0.39132	0.45676	0.01586	0.01531	0.00549	0.00474
gamcytes	0.00006	0.02165	0.87067	0.00972	0.05380	0.00326
fever	0.00292	0.02882	0.05955	0.56580	0.26599	0.00143
newgdclinic	0.00767	0.00497	0.00048	0.00008	0.00043	0.00227

Appendix H. Assessment of confounding

bednet = Survey year (2001 or 1996)
sp2w = SP use (yes/no)
antifolate = Antifolate use (yes/no)
cq2w = CQ use (yes/no)
pfgprev =*pfg377* MS diversity (>1 or ≤1 allele)
newgdshore= GIS distance to lake shore (km)
sex =Sex (male/female)
agemonth = Age (months)
hb =Hemoglobin (g/dL)
lnpmm3 = Parasite density (uL)
gamcytes = Gametocytes present (yes/no)
fever = Fever in last 48 hours (yes/no)
newgdclinic = GIS distance to nearest clinic (km)

Outcome = *dhfr* triple mutant

Model[a]	**Additional variables**	**aOR**	**95% CI Lower**	**95% CI Upper**	**Confounded estimate? (not within 10% of GS aOR)**
1 (GS)	gamcytes, fever, newgdclinic	1.412	0.823	2.423	
2	gamcytes, fever	1.528	0.898	2.598	Not confounded
3	gamcytes, newgdclinic	1.331	0.783	2.264	Not confounded
4	fever, newgdclinic	1.454	0.85	2.487	Not confounded
5	fever	1.572	0.927	2.666	Confounded
6	newgdclinic	1.368	0.806	2.321	Not confounded
7	gamcytes	1.435	0.851	2.419	Not confounded
8*	none	1.475	0.876	2.483	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard
[a] All models contained the variables bednet, antifolate, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described
* Preferred final model

Outcome = *dhfr* triple mutant

Model[a]	Additional variables	aOR	95% CI Lower	95% CI Upper	Confounded estimate? (not within 10% of GS aOR)
1 (GS)	gamcytes, fever, newgdclinic	1.439	0.835	2.478	
2	gamcytes, fever	1.556	0.912	2.655	Not confounded
3	gamcytes, newgdclinic	1.346	0.789	2.298	Not confounded
4	fever, newgdclinic	1.483	0.864	2.546	Not confounded
5	fever	1.604	0.943	2.729	Confounded
6	newgdclinic	1.386	0.814	2.359	Not confounded
7	gamcytes	1.452	0.859	2.456	Not confounded
8*	none	1.498	0.887	2.527	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, sp2w, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Outcome = *dhps* mutations

Model[a]	Additional variables	aOR	95% CI Lower	95% CI Upper	Confounded estimate? (not within 10% of GS aOR)
1 (GS)	gamcytes, fever, newgdclinic	10.362	6.134	17.507	
2	gamcytes, fever	10.214	6.077	17.166	Not confounded
3	gamcytes, newgdclinic	10.419	6.21	17.482	Not confounded
4	fever, newgdclinic	10.237	6.082	17.231	Not confounded
5	fever	10.089	6.026	16.892	Not confounded
6	newgdclinic	10.362	6.189	17.35	Not confounded
7	gamcytes	10.381	6.21	17.354	Not confounded
8*	none	10.319	6.187	17.211	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, antifolate, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Outcome = *dhps* mutations

Model[a]	Additional variables	aOR	95% CI Lower	95% CI Upper	Confounded estimate? (not within 10% of GS aOR)
1 (GS)	gamcytes, fever, newgdclinic	10.472	6.164	17.793	
2	gamcytes, fever	10.347	6.123	17.483	Not confounded
3	gamcytes, newgdclinic	10.537	6.244	17.78	Not confounded
4	fever, newgdclinic	10.398	6.141	17.605	Not confounded
5	fever	10.272	6.101	17.295	Not confounded
6	newgdclinic	10.511	6.241	17.702	Not confounded
7	gamcytes	10.525	6.262	17.69	Not confounded
8*	none	10.496	6.259	17.604	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, sp2w, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Outcome = combined *dhfr/dhps* mutations

Model[a]	Additional variables	aOR	95% CI Lower	95% CI Upper	Confounded estimate? (not within 10% of GS aOR)
1 (GS)	gamcytes, fever, newgdclinic	8.718	5.362	14.174	
2	gamcytes, fever	8.753	5.407	14.169	Not confounded
3	gamcytes, newgdclinic	8.339	5.18	13.425	Not confounded
4	fever, newgdclinic	8.748	5.391	14.196	Not confounded
5	fever	8.786	5.439	14.193	Not confounded
6	newgdclinic	8.379	5.21	13.476	Not confounded
7	gamcytes	8.44	5.259	13.543	Not confounded
8*	none	8.488	5.296	13.605	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, antifolate, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Outcome = combined *dhfr/dhps* mutations

Model[a]	Additional variables	aOR	95% CI Lower	95% CI Upper	Confounded estimate? (not within 10% of GS aOR)
1 (GS)	gamcytes, fever, newgdclinic	8.824	5.404	14.408	
2	gamcytes, fever	8.869	5.459	14.41	Not confounded
3	gamcytes, newgdclinic	8.425	5.212	13.619	Not confounded
4	fever, newgdclinic	8.886	5.453	14.48	Not confounded
5	fever	8.936	5.512	14.486	Not confounded
6	newgdclinic	8.487	5.256	13.704	Not confounded
7	gamcytes	8.542	5.304	13.756	Not confounded
8*	none	8.615	5.357	13.857	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, sp2w, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Outcome = *pfcrt76* mutant

Model[a]	Additional variables	aOR	95% CI Lower	95% CI Upper	Confounded estimate? (not within 10% of GS aOR)
1 (GS)	gamcytes, fever, newgdclinic	1.255	0.73	2.159	
2	gamcytes, fever	1.25	0.732	2.136	Not confounded
3	gamcytes, newgdclinic	1.294	0.759	2.205	Not confounded
4	fever, newgdclinic	1.267	0.739	2.174	Not confounded
5	fever	1.263	0.741	2.151	Not confounded
6	newgdclinic	1.3	0.764	2.213	Not confounded
7	gamcytes	1.27	0.75	2.152	Not confounded
8*	none	1.277	0.755	2.159	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, cq2w, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Outcome = *pfmdr186* mutant

Model[a]	**Additional variables**	**aOR**	**95% CI Lower**	**95% CI Upper**	**Confounded estimate? (not within 10% of GS aOR)**
1 (GS)	gamcytes, fever, newgdclinic	0.797	0.463	1.37	
2	gamcytes, fever	0.8	0.468	1.366	Not confounded
3	gamcytes, newgdclinic	0.773	0.453	1.318	Not confounded
4	fever, newgdclinic	0.789	0.46	1.353	Not confounded
5	fever	0.792	0.465	1.349	Not confounded
6	newgdclinic	0.769	0.452	1.309	Not confounded
7	gamcytes	0.787	0.465	1.334	Not confounded
8*	none	0.783	0.463	1.325	Not confounded

NOTE. aOR, adjusted odds ratio for year of survey variable; CI, confidence interval; GS, gold standard

[a] All models contained the variables bednet, cq2w, lnpmm3, agemonth, sex, hb, pfgprev, and newgdshore in addition to the variables described

* Preferred final model

Appendix I. Assessment of assumption that continuous variables are linear on log scale

In order to verify this assumption, continuous variables were divided into categories (described below). The adjusted odds ratio for the year of survey (bednet variable) in the final model for each outcome was compared between models containing continuous variables and categorical versions of continuous variables. No meaningful differences were observed between adjusted odds ratios for year of survey for all outcomes, therefore we assumed that continuous variables were linearly associated with the log odds of the outcome. Therefore, all final models contained the continuous rather than categorical versions of the variables age, parasite density, hemoglobin, and GIS distance to lake shore.

```
/* agemonth as 3 categories: 0- <=18, 18- <=36, 36+ */
IF agemonth=. THEN agemonthcat=.;
IF agemonth GT 0 AND agemonth LE 18 THEN agemonthcat=0;
IF agemonth GT 18 AND agemonth LE 36.0 THEN agemonthcat=1;
IF agemonth GT 36.0 THEN agemonthcat=2;

/* GDSHORE as 3 categories: 0-5km, 5-10km, 10+km */
IF GDSHORE=. THEN gdshorecat=.;
IF GDSHORE GT 0 AND GDSHORE LE 5000 THEN gdshorecat=0;
IF GDSHORE GT 5000 AND GDSHORE LE 10000 THEN gdshorecat=1;
IF GDSHORE GT 10000 THEN gdshorecat=2;

/* lnpmm3 as 3 categories based on pmm3 */
IF lnpmm3=. THEN lnpmm3cat=.;
IF lnpmm3 GT 0 AND lnpmm3 LE log(2500+1) THEN lnpmm3cat=0;
IF lnpmm3 GT log(2500+1) AND lnpmm3 LE log(10000+1) THEN lnpmm3cat=1;
IF lnpmm3 GT log(10000+1) THEN lnpmm3cat=2;

/* hemoglobin categories */
IF hb=. then hbcat=.;
IF hb LE 7 THEN hbcat=2;
IF hb GT 7 AND hb LT 11 THEN hbcat=1;
IF hb GE 11 THEN hbcat=0;
```

******** CHLOROQUINE RESISTANCE *********

Pfmdr186 Mutant as outcome

```
proc logistic data=genetic_clinic_only_v3;
class lnpmm3cat (ref='0') agemonthcat (ref='2') hbcat (ref='0')
gdshorecat (ref='2') ;
model mdr186=bednet cq2w lnpmm3cat agemonthcat sex hbcat pfgprev
gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	0.803	0.479	1.346
cq2w	1.594	0.802	3.169
lnpmm3cat 1 vs 0	1.308	0.791	2.163
lnpmm3cat 2 vs 0	1.124	0.591	2.137
agemonthcat 0 vs 2	0.425	0.227	0.796
agemonthcat 1 vs 2	0.848	0.477	1.505
sex	1.080	0.701	1.664
hbcat 1 vs 0	1.528	0.857	2.723
hbcat 2 vs 0	1.488	0.673	3.290
pfgprev	1.317	0.845	2.054
gdshorecat 0 vs 2	1.698	0.839	3.438
gdshorecat 1 vs 2	1.293	0.751	2.225

Using continuous variables:

```
proc logistic data=genetic_clinic_only_v3;
model mdr186=bednet cq2w lnpmm3 agemonth sex hb pfgprev newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	0.783	0.463	1.325
cq2w	1.686	0.855	3.321
lnpmm3	1.081	0.950	1.230
agemonth	1.017	1.002	1.032
sex	1.071	0.698	1.644
hb	0.974	0.862	1.100
pfgprev	1.393	0.896	2.167
newgdshore	0.939	0.864	1.02

Pfcrt76 mutant as outcome

```
proc logistic data=genetic_clinic_only_v3;
model crt76=bednet cq2w lnpmm3cat agemonthcat sex hbcat pfgprev
gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	1.217	0.696	2.127
cq2w	1.212	0.593	2.478
lnpmm3cat 1 vs 0	0.864	0.500	1.493
lnpmm3cat 2 vs 0	0.876	0.435	1.763
agemonthcat 0 vs 2	1.015	0.512	2.011
agemonthcat 1 vs 2	1.087	0.591	1.999
sex	1.286	0.794	2.084
hbcat 1 vs 0	0.815	0.415	1.602
hbcat 2 vs 0	1.452	0.553	3.814
pfgprev	0.668	0.407	1.097
gdshorecat 0 vs 2	1.077	0.522	2.225
gdshorecat 1 vs 2	1.407	0.769	2.572

Using continuous variables

```
proc logistic data=genetic_clinic_only_v3;
model crt76=bednet cq2w lnpmm3 agemonth sex hb pfgprev newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	1.299	0.724	2.334
cq2w	1.201	0.591	2.442
lnpmm3	0.975	0.840	1.131
agemonth	0.992	0.976	1.009
sex	1.252	0.774	2.026
hb	1.006	0.875	1.157
pfgprev	0.617	0.378	1.007
newgdshore	0.993	0.906	1.089

********* SULFADOXINE-PYRIMETHAMINE RESISTANCE *********

SP use as drug use variable, dhfr triple mutant as outcome

```
proc logistic data=genetic_clinic_only_v3 descending;
model dhfroutcome=bednet sp2w lnpmm3cat agemonthcat sex hbcat pfgprev
gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	1.371	0.831	2.264
sp2w	0.859	0.307	2.404
lnpmm3cat 1 vs 0	0.795	0.492	1.284
lnpmm3cat 2 vs 0	0.888	0.472	1.671
agemonthcat 0 vs 2	0.609	0.331	1.117
agemonthcat 1 vs 2	0.747	0.431	1.296
sex	0.814	0.533	1.242
hbcat 1 vs 0	1.437	0.798	2.589
hbcat 2 vs 0	1.674	0.767	3.653
pfgprev	2.353	1.514	3.657
gdshorecat 0 vs 2	1.413	0.704	2.835
gdshorecat 1 vs 2	1.096	0.629	1.910

Using continuous variables

```
proc logistic data=genetic_clinic_only_v3 descending;
model dhfroutcome=bednet sp2w lnpmm3 agemonth sex hb pfgprev
newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	1.498	0.887	2.527
sp2w	0.888	0.324	2.433
lnpmm3	1.014	0.891	1.154
agemonth	1.014	1.000	1.029
sex	0.816	0.535	1.244
hb	0.902	0.801	1.016
pfgprev	2.416	1.555	3.752
newgdshore	0.928	0.855	1.008

SP use as drug use variable, dhps mutations as outcome

```
proc logistic data=genetic_clinic_only_v3 ;
class lnpmm3cat (ref='0') agemonthcat (ref='2') hbcat (ref='0')
gdshorecat (ref='2');
model dhpsgen3cat=bednet sp2w lnpmm3cat agemonthcat sex hbcat pfgprev
gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	10.668	6.488	17.541
sp2w	3.054	0.816	11.427
lnpmm3cat 1 vs 0	1.180	0.758	1.838
lnpmm3cat 2 vs 0	1.052	0.588	1.882
agemonthcat 0 vs 2	1.424	0.821	2.470
agemonthcat 1 vs 2	0.939	0.566	1.556
sex	1.099	0.744	1.623
hbcat 1 vs 0	1.278	0.724	2.257
hbcat 2 vs 0	0.978	0.476	2.009
pfgprev	1.034	0.693	1.542
gdshorecat 0 vs 2	1.647	0.887	3.057
gdshorecat 1 vs 2	1.671	1.001	2.787

Using continuous variables:

```
proc logistic data=genetic_clinic_only_v3 ;
model dhpsgen3cat=bednet sp2w lnpmm3 agemonth sex hb pfgprev
newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	10.496	6.259	17.604
sp2w	3.189	0.850	11.962
lnpmm3	1.038	0.916	1.177
agemonth	0.996	0.983	1.010
sex	1.117	0.757	1.648
hb	1.002	0.900	1.115
pfgprev	1.029	0.693	1.527
newgdshore	0.918	0.852	0.989

SP use as drug use variable, combined dhfr/dhps mutations as outcome

```
proc logistic data=genetic_clinic_only_v3 descending;
class lnpmm3cat (ref='0') agemonthcat (ref='2') hbcat (ref='0')
gdshorecat (ref='2');
model genotype=bednet sp2w lnpmm3cat agemonthcat sex hbcat pfgprev
gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	8.374	5.308	13.212
sp2w	1.780	0.671	4.719
lnpmm3cat 1 vs 0	1.021	0.673	1.549
lnpmm3cat 2 vs 0	0.901	0.525	1.545
agemonthcat 0 vs 2	1.373	0.811	2.323
agemonthcat 1 vs 2	0.846	0.531	1.350
sex	0.870	0.604	1.252
hbcat 1 vs 0	1.550	0.931	2.580
hbcat 2 vs 0	1.124	0.567	2.228
pfgprev	1.515	1.043	2.202
gdshorecat 0 vs 2	1.635	0.910	2.939
gdshorecat 1 vs 2	1.629	1.006	2.638

Using continuous variables:

```
proc logistic data=genetic_clinic_only_v3 descending;
model genotype=bednet sp2w lnpmm3 agemonth sex hb pfgprev newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	8.615	5.357	13.857
sp2w	1.887	0.720	4.946
lnpmm3	1.024	0.916	1.146
agemonth	0.998	0.986	1.010
sex	0.896	0.624	1.287
hb	0.972	0.876	1.079
pfgprev	1.518	1.049	2.195
newgdshore	0.908	0.846	0.975

Antifolate use as drug use variable, dhfr triple mutant as outcome

```
proc logistic data=genetic_clinic_only_v3 descending;
class lnpmm3cat (ref='0') agemonthcat (ref='2') hbcat (ref='0')
gdshorecat (ref='2');
model dhfroutcome=bednet antifolate lnpmm3cat agemonthcat sex hbcat
pfgprev gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	1.368	0.832	2.251
antifolate	1.042	0.523	2.078
lnpmm3cat 1 vs 0	0.788	0.490	1.265
lnpmm3cat 2 vs 0	0.861	0.458	1.618
agemonthcat 0 vs 2	0.597	0.327	1.090
agemonthcat 1 vs 2	0.766	0.442	1.325
sex	0.810	0.532	1.231
hbcat 1 vs 0	1.498	0.846	2.653
hbcat 2 vs 0	1.771	0.823	3.809
pfgprev	2.342	1.507	3.641
gdshorecat 0 vs 2	1.456	0.731	2.900
gdshorecat 1 vs 2	1.121	0.648	1.941

Using continuous variables:

```
proc logistic data=genetic_clinic_only_v3 descending;
model dhfroutcome=bednet antifolate lnpmm3 agemonth sex hb pfgprev
newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	1.475	0.876	2.483
antifolate	1.037	0.523	2.056
lnpmm3	1.010	0.888	1.149
agemonth	1.015	1.001	1.030
sex	0.811	0.534	1.232
hb	0.895	0.796	1.006
pfgprev	2.412	1.553	3.747
newgdshore	0.923	0.850	1.002

Antifolate use as drug use variable, dhps mutations as outcome

```
proc logistic data=genetic_clinic_only_v3 ;
class lnpmm3cat (ref='0') agemonthcat (ref='2') hbcat (ref='0')
gdshorecat (ref='2');
model dhpsgen3cat=bednet antifolate lnpmm3cat agemonthcat sex hbcat
pfgprev gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	10.834	6.612	17.751
antifolate	2.310	1.088	4.901
lnpmm3cat 1 vs 0	1.171	0.756	1.813
lnpmm3cat 2 vs 0	1.039	0.582	1.854
agemonthcat 0 vs 2	1.324	0.767	2.286
agemonthcat 1 vs 2	0.932	0.565	1.538
sex	1.048	0.712	1.543
hbcat 1 vs 0	1.404	0.804	2.451
hbcat 2 vs 0	1.086	0.535	2.205
pfgprev	1.078	0.722	1.610
gdshorecat 0 vs 2	1.570	0.849	2.904
gdshorecat 1 vs 2	1.580	0.952	2.624

Using continuous variables:

```
proc logistic data=genetic_clinic_only_v3 ;
model dhpsgen3cat=bednet antifolate lnpmm3 agemonth sex hb pfgprev
newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	10.319	6.187	17.211
antifolate	2.422	1.145	5.126
lnpmm3	1.039	0.916	1.178
agemonth	0.999	0.986	1.012
sex	1.063	0.723	1.562
hb	0.992	0.893	1.103
pfgprev	1.087	0.732	1.613
newgdshore	0.922	0.856	0.993

Antifolate use as drug use variable, combined dhfr/dhps mutations as outcome

```
proc logistic data=genetic_clinic_only_v3 descending;
class lnpmm3cat (ref='0') agemonthcat (ref='2') hbcat (ref='0')
gdshorecat (ref='2');
model genotype=bednet antifolate lnpmm3cat agemonthcat sex hbcat
pfgprev gdshorecat;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	8.450	5.373	13.289
antifolate	1.650	0.881	3.091
lnpmm3cat 1 vs 0	1.012	0.671	1.526
lnpmm3cat 2 vs 0	0.879	0.513	1.505
agemonthcat 0 vs 2	1.312	0.778	2.211
agemonthcat 1 vs 2	0.851	0.535	1.352
sex	0.833	0.580	1.195
hbcat 1 vs 0	1.652	1.003	2.721
hbcat 2 vs 0	1.222	0.623	2.396
pfgprev	1.544	1.062	2.244
gdshorecat 0 vs 2	1.612	0.901	2.884
gdshorecat 1 vs 2	1.584	0.983	2.554

Using continuous variables:

```
proc logistic data=genetic_clinic_only_v3 descending;
model genotype=bednet antifolate lnpmm3 agemonth sex hb pfgprev
newgdshore;
run;
```

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
bednet	8.488	5.296	13.605
antifolate	1.725	0.927	3.208
lnpmm3	1.026	0.917	1.148
agemonth	1.000	0.988	1.012
sex	0.854	0.596	1.223
hb	0.964	0.870	1.069
pfgprev	1.562	1.080	2.258
newgdshore	0.908	0.847	0.974

EMORY UNIVERSITY | Institutional Review Board

TO: Monica Shah
Principal Investigator

DATE: December 9, 2010

RE: **Notification of Submission Determination: No IRB Review Required**
IRB00043231
Effect of transmission reduction by ITNs on the prevalence of mutations associated with resistance to sulfadoxine-pyrimethamine and chloroquine in western Kenya

The above-referenced study has been vetted by the Institutional Review Board (IRB), and it was determined that it does not require IRB review because it does not meet the definition of "Research involving Human Subjects" under applicable federal regulations. Based on the information included in the submission, this proposed secondary analysis will investigate the effect of sustained transmission reduction by ITNs on the prevalence of genes associated with resistance to the anti-malarial drugs SP and CQ in children under five years old. The PI will use existing clinical, epidemiological and genetic data collected during an ITN trial conducted by the CDC and KEMRI. This dataset has been de-identified and the PI will not have access to codes linking identifiers to the participants now or in the future. Accordingly, IRB review is not required.

45 CFR Section 46.102(f)(2) defines "Research involving Human Subjects" as follows:

> Human subject means a living individual about whom an investigator (whether professional or student) conducting research obtains:
>
> (1) data through intervention or interaction with the individual, or
> (2) identifiable private information
>
> Intervention includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. Interaction includes communication or interpersonal contact between investigator and subject. Private information includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). Private information must be individually identifiable (i.e., the identity of the subject is or may be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects.

Please note that any changes to the protocol could conceivably alter the status of this research under the federal regulations cited above. Accordingly, any substantive changes in the protocol should be presented to the IRB for consideration prior to their implementation in the research.

Sincerely,

Carol Corkran, MPH, CIP
Senior Research Protocol Analyst
This letter has been digitally signed