```
## Error in library(ggmap):  there is no package called 'ggmap'
```

1 Abstract

Malaria is a significant threat to public health in countries where the disease is either endemic or epidemic. Concerted efforts have been made in the past decade to reduce and in some cases eliminate malaria with the use of prophylactic interventions. The World Health Organization recommends preferential administration of interventions to pregnant women and infants because of the high disease burden borne by this group. However, previous research has identified the benefit of additionally targeting interventions at those with the highest risk of infections. We use a topographic wetness index, combined with a household census of intervention use, at two sites in Kenya to assess intervention administration. We find preferential administration of interventions at the high-elevation epidemic-prone site but not at the low-elevation endemic site. Our results have important implications for assessing the administration of interventions in the battle against malaria.

2 Introduction

Malaria and other mosquito borne illnesses are considered a significant threat to public health and a socio-economic burden in countries where these diseases are either endemic or epidemic [?]. Concerted efforts have been made in the past decade to reduce and in some cases eliminate malaria specifically. Many national strategic plans to reduce or eliminate malaria are in their third generation. Spatial targeting of high risk areas is a strategy that has been recommended but few studies have assessed if government programs are achieving differential coverage in high risk areas [?].

The government of Kenya developed the National Malaria Strategy 2009-2017 in response to the ongoing threat of malaria [?]. This strategy outlined 6 objectives, the first of which is to have at least 80% of people living in malaria risk areas using appropriate malaria preventive interventions. The two primary non-pharmaceutical interventions identified in the plan are Indoor Residual Spraying (IRS) and Long Lasting Insecticidal Nets (LLINs). The strategy outlined for achieving the intervention objective included the initial mass distribution of LLINs where malaria is either endemic (western lowlands) or epidemic-prone (western highlands); followed by routine distribution of LLINs to pregnant women and children under 1 year of age and a subsidized sale of LLINs. The strategy also outlined the use of widespread IRS followed by focal treatments in epidemic-prone areas.

The World Health Organization recommends prioritizing the administration of interventions to pregnant women and young children followed by progressively achieving intervention coverage of all community members. The preferential administration of interventions to pregnant women and young children reflects the disproportionate disease burden borne by this group [?]. However, previous research has identified the benefit of additionally targeting interventions at those with the highest risk of infections [?]. Identifying individuals which are both vulnerable to infection and likely to be exposed to an infected mosquito is therefore a priority. However, exisiting distribution campains do not typically account for both infection risk and disease burden simulataneously [?].

Remotely-sensed topographic data has been previously investigated as a tool for assessing risk of malaria infection by identifying areas where water is likely to pool and Anopheles densities are likely to be higher [?, ?]. This method, therefore, has the potential both inform new distribution campains and evaluate the efficacy of existing campains. Our primary objective was to use topographic data, combined with a household census of demographics and intervention use, to develop a method for identifying high risk households. We applied our method to two sites in Kenya where malaria is either endemic or epidemic-prone. Since policies for intervention administration differed between the epidemic-prone and endemic regions we also sought to compare the efficacy of intervention distribution between these two regions.

3 Statistical Methodology

3.1 Topographical Wetness Index

The Topographical Wetness index was originally introduced by Bevin and Kirby in 1979 [?]. TWI combines a measure of the amount of upstream drainage area with the local slope to determine the amount of wetness likely to accumulate at a point and is defined as:

$$ln$$

$$\tan b,$$

$$where$$

$a is the local upslope area and \tan b is the local slope in radians. Since TWI was first defined numerous methods have b$
$flow. We did this in two steps. We first identified local depressions by determining, for each cell in the study region,$

We utilized 90 meter resolution elevation data from the National Aeronautics and Space Administration (NASA) Shuttle Radar Topography Mission (SRTM). The high elevation site was sufficiently covered by tile number 43-12 but we utilized two adjacent tiles (43-13 and 44-13) in order to eliminate possible edge effects for eastern households at the low site.

4 Description of the Data

4.1 House Hold Data

Prior to the initiation of a community-based research program, two study sites in western Kenya were mapped and a census was taken. These two sites represent the western highland (hereafter epidemic-prone, N = 3380) and lowland (hereafter endemic, N = 604) populations. We collected demographic information for each occupant including age and sex. Both sites have had partial treatment with both LLINs and IRS and household heads provided initial information about LLIN ownership and government administration of household IRS in the previous six months. Additional information for each participant was also collected such as age, sex, and relation to the head of the household.

```
## Error in library(BDSS): there is no package called 'BDSS'
```

4.2 Geo-Spatial Data

5 Study Objectives

5.1 Primary Objective

Our primary objective was to determine whether populations at high combined risk for both exposure to, and a poor outcome from, malaria are receiving mitigating treatments at a higher rate than those with a low combined risk. Specifically we will test the following alternative hypotheses:

H_0 : High-risk households are not more likely to receive treatment than low-risk households. H_a : High-risk households are more likely to receive treatment than low-risk households.

5.2 Secondary Objectives

As a secondary objective, we will determine if populations which are only at risk for a poor health outcome from malaria are preferentially receiving mitigating treatments. By combining measures from the primary and secondary outcomes we will determine if there would be a benefit from modifying treatment administration protocols to incorporate information on the risk of mosquito exposure in addition to the current strategy. We will also determine if individuals at risk due to old age are more likely to be missed by current protocols.

6 Statistical Methods

The unit of study for this analysis is the household since both bed nets and aerial spraying are administered at the household level. Therefore, we summarized the

information from the individual surveys into household attributes. We identified unique households by the unique combination of sub-location, village, and house number. We found and eliminated duplicate individual surveys by identifying entries with identical house identification and age attributes. This potentially removed non-duplicate individuals (such as twins in the same household) but we felt the introduced bias would be negligible. For each unique house we calculated the number of individuals under 1, the number of individuals over 1 and under 5, and the number of individuals over 65. We also determined if each house had received a bed net or aerial spraying. Since responses among household members was not consistent we assigned a treatment to the house if any member of the house responded affirmatively.

Since the two sites have substantially different rates of spraying and bed net usage we will analyze the high and low sites separately with regards to these outcomes. We also analyzed spraying and bed net usage separately within each site since these are known to be distributed to households under different protocols and are therefore likely to have different patterns.

Each household varied with respect to both the risk of exposure to mosquitoes and the number of at risk individuals in the household. We assigned an age-based health risk score (age-based risk hereafter) to each household with the following formula:

$$RiskScore = (2 \times Children \leq 1) + (1 < Children \leq 5) + (Adults > 65)$$

We assigned twice the weight to children under 1 since they have the highest risk of the categories (Gupta et al. 1999 and Snow et al. 1999). We did not have information on pregnancy so we could not include this in the current risk assessment although there is a known risk for premature birth (Menendez et al. 2000).

We assigned each household a risk for exposure to mosquitoes (mosquito-based risk hereafter) by deriving a continuous risk surface over the study area. We used a Topographical Wetness Index (TWI) derived from the DEM data to determine areas likely to provide breeding habitat for mosquitoes. The TWI combines the total basin area (the area from which water will flow to a particular point) with the slope at that point to determine the amount of water likely to accumulate and provide breeding habitat for mosquitoes. We used the TWI algorithm provided by the open source System for Automated Geo-scientific Analyses (SAGA) to locate areas of high wetness. We assumed the mosquito exposure risk of a household was inversely related to the distance to one or more of these high-wetness areas. Therefore, we applied a Gaussian filter with $\sigma = 10$ to create a weighted average of mosquito risk for each cell in the study area. We then assigned each house the risk score of the cell it was in.

6.1 Primary Analysis

To be at risk for a poor outcome a person must 1) come in contact with a malaria hosting mosquito, and 2) be inherently vulnerable to malaria infection

4

(i.e. very young or very old). We will create two risk scores representing each of these household risks. The household health risk will be equal to the number of individuals at high risk (i.e. under age 5 or over age 65). The household exposure risk will be equal to the score of the risk surface at the house location. Since these risks will be calculated on different scales we will center at 0 and standardize risk scores so that they are scale-independent.

We will add the standardized household health risk with the standardized household exposure risk to create a combined risk. We will then determine if high risk households are more likely to have received either a bed-net or aerial spraying with a logistic model;

$$log(\frac{p}{1-p}) = \beta_0 + \beta_1 \times CombinedHouseholdRisk,$$

where p = Probability of a house having a treatment. If e^{β_1} is ¿ 1 and statistically significant ($\alpha = 0.05$) then high-risk households are more likely to receive treatment. We used a restricted cubic spline function to determine if there was a linear relationship between the log odds of treatment and combined risk. If we found evidence of a non-linear relationship we categorized the risk score into quartiles and re-fit with a means model.

6.2   Secondary Analyses

Existing protocols may be adequate at addressing household risk due to either inherent health risk or mosquito exposure risk, but not both. We will conduct the same primary analysis but separate out these two risk scores as separate predictors:

$$log(\frac{p}{1-p}) = \beta_0 + \beta_2 \times MosquitoExposureRisk + \beta_3 \times HealthRisk,$$

where p = Probability of a house having a treatment. The interpretation of β_2 and β_3 is the same as β_1 from the primary analysis, but specific to a risk type.

6.3   Sensitivity Analyses

Bed nets are currently targeted at pregnant women so we expect that bed net usage will be higher in households with young children. However, our age-based risk score also incorporates elderly household members. Therefore, in order to determine if current protocols are effective in targeting pregnant women and young children, we will re-define the age based risk to only include the two young child age categories. We will repeat the above analysis with the restricted age-based risk score for both bed nets and aerial spraying.

```
## Error in library(SDMTools):  there is no package called 'SDMTools'
## Error in eval(expr, envir, enclos):  could not find function "aspect"
## Error in eval(expr, envir, enclos):  object 'lowasp' not found
## Error in values(lowasp):  error in evaluating the argument 'x' in
selecting a method for function 'values':  Error:  object 'lowasp'
not found
## Error in focal(lowaspc, aspmat3, fun = var, na.rm = T): error in
evaluating the argument 'x' in selecting a method for function 'focal':
Error:  object 'lowaspc' not found
## Error in focal(lowaspc, aspmat7, fun = var, na.rm = T): error in
evaluating the argument 'x' in selecting a method for function 'focal':
Error:  object 'lowaspc' not found
## Error in focal(lowaspc, aspmat11, fun = var, na.rm = T): error in
evaluating the argument 'x' in selecting a method for function 'focal':
Error:  object 'lowaspc' not found
## Error in eval(expr, envir, enclos):  object 'aspvar3' not found
## Error in eval(expr, envir, enclos):  object 'lowaspvar' not found
## Error in eval(expr, envir, enclos):  could not find function "aspect"
## Error in eval(expr, envir, enclos):  object 'highasp' not found
## Error in values(highasp):  error in evaluating the argument 'x'
in selecting a method for function 'values':  Error:  object 'highasp'
not found
## Error in focal(highaspc, aspmat3, fun = var, na.rm = T): error in
evaluating the argument 'x' in selecting a method for function 'focal':
Error:  object 'highaspc' not found
## Error in focal(highaspc, aspmat7, fun = var, na.rm = T): error in
evaluating the argument 'x' in selecting a method for function 'focal':
Error:  object 'highaspc' not found
## Error in focal(highaspc, aspmat11, fun = var, na.rm = T): error
in evaluating the argument 'x' in selecting a method for function 'focal':
Error:  object 'highaspc' not found
## Error in eval(expr, envir, enclos):  object 'aspvar3' not found
## Error in eval(expr, envir, enclos):  object 'highaspvar' not found
## Error in as.data.frame(hightwi, xy = TRUE): error in evaluating
the argument 'x' in selecting a method for function 'as.data.frame':
Error:  object 'hightwi' not found
## Error in as.data.frame(lowtwi, xy = TRUE): error in evaluating the
argument 'x' in selecting a method for function 'as.data.frame':  Error:
object 'lowtwi' not found
```

7 Results

The odds of receiving either a bed net or aerial spraying are higher for households with higher combined risk, but only at the high site (table 1). For each 1

standard deviation increase in combined risk at the high site the probability of receiving a bed net increases 27

The probability of bed net use at the high site was more strongly associated with age-based risk, whereas the probability of aerial spraying at the high site was more strongly associated with mosquito-based risk (table 3). However, We did not find the same pattern at the low site where we found households with high mosquito-based risk were actually significantly less likely to receive aerial spraying (OR: 0.35, 95

Table 1. Odds of receiving a treatment as a function of combined risk.

Site	Treatment	OR	Lower 95
High	Net	1.35	1.26 1.44
	Spray	1.13	1 1.26
Low	Net	1.09	0.86 1.39
	Spray	0.9	0.61 1.34

Table 2. Odds of treatment from risk of either mosquito exposure or malaria risk.

Site	Treatment	Age Risk	Age Risk	Age Risk	Mosquito Risk	Mosquito Risk	Mosquito Risk
		OR	Lower	95			
High	Net	1.36	1.27	1.45	1.01	0.93	1.1
	Spray	1.08	0.96	1.22	1.32	1.14	1.53
Low	Net	1.21	0.94	1.55	0.58	0.31	1.1
	Spray	1.08	0.67	1.74	0.34	0.15	0.79

```
## Error in eval(expr, envir, enclos):  object 'lowtwi' not found
## Error in eval(expr, envir, enclos):  object 'hightwi' not found
## Error in focal(lowtr_s, fmat, filename = "smoothLowTWI_sens", pad
= T, :  error in evaluating the argument 'x' in selecting a method
for function 'focal':  Error:  object 'lowtr_s' not found
## Error in focal(hightr_s, fmat, filename = "smoothHighTWI_sens", pad
= T, :  error in evaluating the argument 'x' in selecting a method
for function 'focal':  Error:  object 'hightr_s' not found
```

```
## Error in extract(mriskLows, y = cbind(lowsite$lon, lowsite$lat)):
error in evaluating the argument 'x' in selecting a method for function
'extract':  Error:  object 'mriskLows' not found
## Error in extract(mriskHighs, y = cbind(highsite$lon, highsite$lat)):
error in evaluating the argument 'x' in selecting a method for function
'extract':  Error:  object 'mriskHighs' not found
```

```
## Error in scale(log(lowsite$mrisks + 1), center = F) + scale(log(lowsite$arisk
+ :  non-conformable arrays
## Error in scale(log(highsite$mrisks + 1), center = F) + scale(log(highsite$arisk
+ :  non-conformable arrays
```

```
## Error in eval(expr, envir, enclos):  object 'comrsks' not found
## Error in eval(expr, envir, enclos):  object 'comrsks' not found
```

```
## Error in eval(expr, envir, enclos):  object 'comrsks' not found
## Error in eval(expr, envir, enclos):  object 'comrsks' not found
## Error in eval(expr, envir, enclos):  object 'comrsks' not found
## Error in eval(expr, envir, enclos):  object 'comrsks' not found
```

7.1 Sensitivity Analysis

The use of the restricted TWI algorithm identified fewer regions as high risk than the SAGA packaged algorithm at both sites (fig. 2). The use of the restricted TWI based risk surface in the combined risk score increased the odds of high-risk households receiving a treatment for both the high and low sites, although the increase in OR for the low site remained non-significant (table 3).

Eliminating elderly household members from the age-based risk calculation increased the odds ratio for net use at both the high and low sites (Table 4). However, the OR for aerial spraying decreased slightly at both sites.

Table 3. Comparison of restricted TWI results with general TWI results.

Site	Treatment	OR	Lower	95			
High	Net	1.35	1.26	1.44	1.35	1.26	1.44
	Spray	1.13	1	1.26	1.13	1	1.26
Low	Net	1.09	0.86	1.39	1.09	0.86	1.39
	Spray	0.9	0.61	1.34	0.9	0.61	1.34

Table 4. Comparison of the odds of receiving a treatment based on health risk due to age with and without the inclusion of elderly household members.

Site	Treatment	Age Risk (¡5 or ¿65)	Age Risk (¡5 or ¿65)	Age Risk (¡5 or ¿65)	Age Risk (¡5)	Age Risk (¡5)	Age Risk (¡5)
		OR	Lower	95			
High	Net	1.36	1.27	1.45	1.36	1.27	1.45
	Spray	1.08	0.96	1.22	1.08	0.96	1.22
Low	Net	1.21	0.94	1.55	1.21	0.94	1.55
	Spray	1.08	0.67	1.74	1.08	0.67	1.74

8   Discussion

Current protocols for administration of bed nets target pregnant women. Therefore, we would expect that households with young children would be more likely to have bed nets. We found that age-based risk was associated with an increased probability of bed net use at the high site but not the low site. Although the association of bed net use with high age-based risk improved slightly when elderly adults were removed from the risk calculation, it was still not substantial or significant.

Aerial spraying is intended to target households at high risk for mosquito exposure under current protocols so we would expect that households with high mosquito-based risk would be associated with aerial spraying. Again, this is the pattern we observed at the high site but not the low site where we found the opposite association. Use of the restricted TWI algorithm suggested that the association was at least in the preferable direction at the low site but not significantly. The sensitivity of our results to choice of TWI algorithm suggests that the TWI should be validated with additional information such as ground-truthing or infection rate data. This has been done previously (CITE CITE), but only with a single algorithm.