1 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 (*The Second National Health Sector Strategic Plan of Kenya: NHSSP II 20052010*). 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 (Schantz-Dunn and Nour 2009).

The government of Kenya developed the National Malaria Strategy 2009-2017 in response to the ongoing threat of malaria (health and sanitation 2009). 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 Bousema et al. 2012. However, previous research has identified the benefit of additionally targeting interventions at those with the highest risk of infections Schantz-Dunn and Nour 2009. 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.

Freely available 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 Cohen et al. 2008; Cohen et al. 2010. This method, therefore, has the potential to 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. Becuase of the wide variety of TWI algorithms used in practice we further wanted to investigate the sensitivity of the results to the choice of TWI algorithm. We applied the method to two sites in Kenya where malaria is either endemic or epidemic-prone. Since policies for intervention administration, as well as topographic features, differed between the epidemic-prone and endemic regions we also sought to compare the sensitivity of the method to choice of TWI algorithm between these two regions.

2 Statistical Methodology

Every household on a landscape will vary with respect to both the risk of exposure to mosquitoes and the number of at risk individuals in the household. We develop a method to combine these two risk factors into an overall risk score for each household within a managment area in order to identify the households which will yeild the greatest benefit from the application of limited profalactic resources.

2.1 Topographical Wetness Index

The Topographical Wetness index was originally introduced by Beven and Kirby in 1979 (BEVEN and KIRKBY 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 by Beven and Kirby as:

$$ln\frac{a}{tanb},$$

where a is the local upslope area and $tanb$ is the local slope in radians. The TWI is designed to predict the amount of water that is likely to flow into an area, based on surface topology, and the rate at which this water will flow out of an area. Areas with high in-flow and low out-flow are likely provide habitat for breeding mosquitoes.

Since TWI was first defined numerous methods have been developed to calculate it and several review papers have been published Quinn, Beven, and Lamb 1995; Sørensen, Zinko, and Seibert 2006 as well as alternative modeling strategies developed Grabs et al. 2009. The primary goal of using TWI in this application is to identify areas where mosquitoes are likely to breed. However, it is unclear what effect differing TWI algorithms would have on our results. To determine the sensitivity of our method to choice of TWI algorithm, we implement 3 diffent algorithms which differ in their calculation of the upslope area and local slope: 1) traditional TWI BEVEN and KIRKBY 1979, 2) the SAGA Wetness Index Böhner 2002, and 3) a simplified terrain wetness index which simply identifies depressions. We carried out all analyses using the statistical programming language R version 3.2.3 R Core Team 2015, with the exception of the SAGA TWI which was calculated with the SAGA open-source GIS software Böhner and Selige 2006. Details for the calculation of all three methods are provided in appendix A.

We assign each household a risk for exposure to mosquitoes (mosquito-based risk hereafter) by deriving a continuous risk surface over the study area from each TWI algorithm. We assume the mosquito exposure risk of a household is inversely related to the distance to one or more of these high-wetness areas. Therefore, we apply a Gaussian filter with $\sigma = 10$ to create a weighted average of mosquito risk for each cell in the study area. We then assign each house the risk score of the cell it is in or, for high resolution topographical data such as LiDAR, the average of the cells a property occupies.

2.2 Individual Health Risk

Each household varied with respect to both the risk of exposure to mosquitoes and the number of at risk individuals in the household. The household risk formulae will vary depending on the disease under study and the most vulnerable population(s) for that disease. We recommend a simple additive risk score, like the score formula provided below developed for malaria, based on expert opinion or relevant literature. The sole purpose of the health risk score is to differentiate households with high risk from those with low risk. More complicated formulae can be constructed but we beleive a simple formula will be easier to interpret and adequate for most applications.

2.3 Overall Risk

To be at risk for a poor outcome a person must 1) come in contact with a disease harboring mosquito, and 2) be inherently vulnerable to infection (e.g. very young or very old). We create two risk scores representing each of these types of risk for every household on the landscape. Since these risks will be calculated on different scales we will center at 0 and standardize risk scores so that they are scale-independent. These two scores can then be combined with a weighted sum to create an overall risk score for the household. Weights in the sum are determined by expert opinion, in our example below we use equal weights. This method lends itself well to the addition of additional risk scores, such as the distance to a health care provider, or even risk attenuating factors such as the presence of window screens, which can all be summed together with appropriate weights.

2.4 Example Analysis

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. Our objective was to determine if the highest risk households were given preference in the administration of interventions in the form of both bed nets and indoor residual spraying.

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.

We assigned an individual-based health risk score (age-based risk hereafter) to each household with the following formula:

$$Risk\ Score = (2 \times\ Children\ \leq 1) + (1 <\ Children\ \leq 5) + (2 \times\ Pregnant\ Women)$$

We assigned twice the weight to children under 1 and pregnant women since they have been previously identified as high risk Gupta et al. 1999; Snow et al. 1999; Menendez, Fleming, and Alonso 2000. This formula is tailored for malaria exposure but can be adjusted based on *a priori* information about the individual risks for a particular mosquito borne disease. For example, for the Zica virus children may have a relatively low health risk whereas pregnant women and sexually active women may have much higher risk.

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

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

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.

In order to determin the sensitivity of this method to the choice of TWI algorithm, we repeated the analysis using each of the three TWI algorithms described in appendix A.

```
## Warning:  package 'raster' was built under R version 3.2.3
## Warning:  package 'sp' was built under R version 3.2.3
## Warning:  package 'rgdal' was built under R version 3.2.3
```

2.5 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% (OR: 1.27, 95% CI: 1.18, 1.35) and the probability of aerial spraying increases 15% (OR: 1.15, 95% CI: 1.03, 1.29). At the low site, we found no preferential administration of either treatment to high combined risk households. We found some evidence, from the fitting of a restricted cubic spline, of a non-linear relationship between the log-odds of net use and combined risk at the low site. However, modelling the mean risk for each risk quantile did not change our results.

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% CI: 0.14, 0.83).

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

	Site	Treatment	OR	Lower 95% CI	Upper 95% CI
1	High	Net	1.35	1.26	1.44
2		Spray	1.13	1	1.26
3	Low	Net	1.09	0.86	1.39
4		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
1			OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI
2	High	Net	1.36	1.27	1.45	1.01	0.93	1.1
3		Spray	1.08	0.96	1.22	1.32	1.14	1.53
4	Low	Net	1.21	0.94	1.55	0.58	0.31	1.1
5		Spray	1.08	0.67	1.74	0.34	0.15	0.79

2.6 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% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI
1	High	Net	1.35	1.26	1.44	1.45	1.33	1.59
2		Spray	1.13	1	1.26	1.22	1.05	1.42
3	Low	Net	1.09	0.86	1.39	1.33	0.93	1.9
4		Spray	0.9	0.61	1.34	1.1	0.59	2.03

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.

3 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

	Site	Treatment	Age Risk (<5 or >65)	Age Risk (<5 or >65)	Age Risk (<5 or >65)	Age Risk (<5)	Age R
1			OR	Lower 95% CI	Upper 95% CI	OR	Lower
2	High	Net	1.36	1.27	1.45	1.36	1.27
3		Spray	1.08	0.96	1.22	1.08	0.96
4	Low	Net	1.21	0.94	1.55	1.21	0.94
5		Spray	1.08	0.67	1.74	1.08	0.67

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 (Cohen2008, Cohen2010), but only with a single algorithm.

References

BEVEN, K. J. and M. J. KIRKBY (1979). "A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant". en. In: *Hydrological Sciences Bulletin* 24.1, pp. 43–69. ISSN: 0303-6936. DOI: 10.1080/02626667909491834. URL: http://www.tandfonline.com/doi/abs/10.1080/02626667909491834.

Böhner, J (2002). *Soil Regionalisation by Means of Terrain Analysis and Process Parameterisation.* Ispra.

Böhner, J and T Selige (2006). "Spatial prediction of soil attributes using terrain analysis and climate regionalisation". In: *Gottinger Geographische Abhandlungen.* URL: http://www.pe.wzw.tum.de/publikationen/pdf/sd663.pdf.

Bousema, Teun et al. (2012). "Hitting hotspots: spatial targeting of malaria for control and elimination." In: *PLoS medicine* 9.1, e1001165. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.1001165. URL: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001165.

Cohen, Justin M et al. (2008). "Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western Kenyan highlands." En. In: *Malaria journal* 7.1, p. 40. ISSN: 1475-2875. DOI: 10.1186/1475-2875-7-40. URL: http://malariajournal.biomedcentral.com/articles/10.1186/1475-2875-7-40.

— (2010). "Local topographic wetness indices predict household malaria risk better than land-use and land-cover in the western Kenya highlands." In: *Malaria journal* 9, p. 328. ISSN: 1475-2875. DOI: 10.1186/1475-2875-9-328. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2993734{\&}tool=pmcentrez{\&}rendertype=abstract.

Crouch, Margaret. *The Second National Health Sector Strategic Plan of Kenya: NHSSP II 20052010.* Tech. rep. Republic of Kenya Ministry of Health. URL: http://www.healthpolicyinitiative.com/Publications/Documents/796{_}1{_}796{_}1{_}NHSSP{_}II.pdf.

Grabs, T. et al. (2009). "Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model". In: *Journal of Hydrology* 373.1-2, pp. 15–23. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2009.03.031. URL: http://www.sciencedirect.com/science/article/pii/S002216940900211X.

Gupta, S et al. (1999). "Immunity to non-cerebral severe malaria is acquired after one or two infections." In: *Nature medicine* 5.3, pp. 340–3. ISSN: 1078-8956. DOI: 10.1038/6560. URL: http://www.ncbi.nlm.nih.gov/pubmed/10086393.

health, Ministry of public and sanitation (2009). *National Malaria Strategy 2009-2017*. Tech. rep. July. Ministry of public health and sanitation, p. 116. URL: https://www.google.com/url?sa=t{\&}rct=j{\&}q={\&}esrc=s{\&}source=web{\&}cd=1{\&}ved=0CB4QFjAAahUKEwiXj76j-IzHAhWGVT4KHZD7AFA{\&}url=http://www.c-hubonline.org/sites/default/files/resources/main/Kenya{_}National{_}Malaria{_}Strategy{_}2009-2017.pdf{\&}ei=9GG{_}VZfFKIar-QGQ94OABQ{\&}usg=AFQj.

Menendez, C, A F Fleming, and P L Alonso (2000). "Malaria-related anaemia." In: *Parasitology today (Personal ed.)* 16.11, pp. 469–76. ISSN: 0169-4758. URL: http://www.ncbi.nlm.nih.gov/pubmed/11063857.

Quinn, P. F., K. J. Beven, and R. Lamb (1995). "The in(a/tan/β) index: How to calculate it and how to use it within the topmodel framework". In: *Hydrological Processes* 9.2, pp. 161–182. ISSN: 08856087. DOI: 10.1002/hyp.3360090204. URL: http://doi.wiley.com/10.1002/hyp.3360090204.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. URL: https://www.r-project.org/.

Schantz-Dunn, Julianna and Nawal M Nour (2009). "Malaria and pregnancy: a global health perspective." In: *Reviews in obstetrics & gynecology* 2.3, pp. 186–92. ISSN: 1941-2797. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2760896{\&}tool=pmcentrez{\&}rendertype=abstract.

Snow, R W et al. (1999). "Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population." In: *Bulletin of the World Health Organization* 77.8, pp. 624–40. ISSN: 0042-9686. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2557714{\&}tool=pmcentrez{\&}rendertype=abstract.

Sørensen, R., U. Zinko, and J. Seibert (2006). "On the calculation of the topographic wetness index: evaluation of different methods based on field observations". English. In: *Hydrology and Earth System Sciences* 10.1, pp. 101–112. ISSN: 1607-7938. DOI: 10.5194/hess-10-101-2006. URL: http://www.hydrol-earth-syst-sci.net/10/101/2006/hess-10-101-2006.html.

4 Appendix A