Malaria Prevention: Are high-risk households receiving treatments?

Dominic D. LaRoche and Dr. Kacey Ernst

Friday, December 05, 2014

# Introduction

Mosquitoes are a problematic transmission vector for a number of infectious diseases in tropical and sub-tropical regions throughout the globe. These diseases, such as malaria, can be particularly deadly for vulnerable populations with restricted access to healthcare. Several promising methods for mitigating the risk of infection from mosquitoes have been deployed in recent years in high risk areas. The Center for Disease Control (CDC) has been actively purchasing bed nets, and working with local organizations in Kenya to distribute them, since 2006. Bed nets reduce the number of encounters with mosquitoes and the diseases they harbor. Another method for reducing the risk of mosquito born illnesses is the use of aerial pesticides to reduce the local mosquito population. Both of these methods have been employed in Kenya in recent years.

Both bed nets and aerial spraying take time and money to deploy. Therefore, it is critical that treatments are first applied to populations at the highest risk of exposure to malaria and only later applied to populations with low risk of exposure. Some measures have been taken to ensure at risk populations receive priority when administering treatment, such as distributing bed nets to pregnant women or spraying households at high risk for mosquito encounters. The World Health Organization (WHO) currently recommends making children and pregnant women the highest priority for mosquito abatement treatments. However, it is unclear whether these measures are adequate in prioritizing mitigation for at risk populations. The purpose of this study is to determine whether individuals with the highest risk of both exposure and poor outcome from malaria infection are more likely to receive a mitigating treatment.

# Data Description

## House Hold Data

We use individual survey data of 17,823 members of 3,924 households at two sites in Kenya. These two sites represent high elevation and low elevation populations. Both sites have had partial treatment with both bed nets and aerial spraying. The high site has more prevalent bed net usage whereas the low site has more prevalent aerial spraying. Each occupant of the 3,924 households was interviewed about bed-net usage and household spraying. Additional information for each participant was also collected such as age, sex, and relation to the head of the household.

## Geo-Spatial Data

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.

# Study Objectives

## 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 hypotheses:

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

## 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.

# 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 were 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:

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 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.

## 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 (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;

where p = Probability of a house having a treatment. If is > 1 and statistically significant () 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.

## 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:

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

## 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 which may inhibit our ability to assess the efficacy of current protocols. Therefore, in order to determine if current protocols are effective in 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.

There are currently many ways to calculate TWI in practice and it is unclear what effect a different TWI algorithm would have on our results. In order to determine the sensitivity of our results to choice of TWI algorithm we calculated a second mosquito risk surface with tighter restrictions on water out-flow. We did this in two steps. We first identified local depressions by determining, for each cell in the study region, if the cell had a lower elevation than the mean of its neighboring cells. This will effectively identify valleys but not necessarily pools, i.e. where water outflow is likely to be low. Therefore, we also calculated the aspect variance of the neighboring cells. Areas with high aspect variance are likely to form pools or peaks. Combining the two measures will identify only depressions with low water outflow. We repeated the analyses described above with the restricted TWI algorithm.

## 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% (Odds Ratio (OR): 1.26, 95% Confidence Interval (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 2). However, we did not find the same pattern at the low site where we found an increased mosquito-based risk was actually significantly associated with a lower odds of receiving aerial spraying (OR: 0.35, 95% CI: 0.14, 0.83).

Table 1. Odds of receiving a treatment as a function of combined risk. Odds ratios represent the effect of an increase of 1 standard deviation in the combined risk measure.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Site | Treatment | OR | Lower 95% CI | Upper 95% CI |
| High | Net | 1.26 | 1.18 | 1.35 |
|  | Spray | 1.15 | 1.03 | 1.29 |
| Low | Net | 0.95 | 0.73 | 1.24 |
|  | Spray | 0.98 | 0.64 | 1.50 |

Table 2. Odds of treatment from risk of either mosquito exposure or malaria risk. Odds ratios represent the effect of an increase of 1 standard deviation in either risk measure.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Age-Based Risk | | | Mosquito-Based Risk | | |
| Site | Treatment | OR | 95% Confidence Interval | | OR | 95% Confidence Interval | |
| High | Net | 1.27 | 1.18 | 1.35 | 1.01 | 0.93 | 1.1 |
|  | Spray | 1.11 | 0.99 | 1.25 | 1.32 | 1.14 | 1.53 |
| Low | Net | 1.05 | 0.8 | 1.38 | 0.58 | 0.31 | 1.1 |
|  | Spray | 1.22 | 0.73 | 2.05 | 0.34 | 0.15 | 0.79 |

## Sensitivity Analysis

The use of the restricted TWI algorithm identified fewer regions as high risk than the SAGA packaged algorithm at both sites (fig. 1). 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. Odds ratios reflect an increase of 1 standard deviation in risk based on either TWI.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | SAGA TWI | | | Restricted TWI | | |
| Site | Treatment | OR | 95% Confidence Interval | | OR | 95% Confidence Interval | |
| High | Net | 1.26 | 1.18 | 1.35 | 1.34 | 1.22 | 1.47 |
|  | Spray | 1.15 | 1.03 | 1.29 | 1.33 | 1.12 | 1.57 |
| Low | Net | 0.95 | 0.73 | 1.24 | 1.07 | 0.69 | 1.64 |
|  | Spray | 0.98 | 0.64 | 1.5 | 1.31 | 0.63 | 2.69 |

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. Odds ratios reflect an increase of 1 standard deviation in age-based risk.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Age Risk (<5 or >65) | | | Age Risk (<5) | | |
| Site | Treatment | OR | 95% Confidence Interval | | OR | 95% Confidence Interval | |
| High | Net | 1.27 | 1.18 | 1.35 | 1.36 | 1.27 | 1.45 |
|  | Spray | 1.11 | 0.99 | 1.25 | 1.08 | 0.96 | 1.22 |
| Low | Net | 1.05 | 0.8 | 1.38 | 1.21 | 0.94 | 1.55 |
|  | Spray | 1.22 | 0.73 | 2.05 | 1.08 | 0.67 | 1.74 |

# 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 (Cohen et al. 2008 and 2010), but only with a single algorithm.

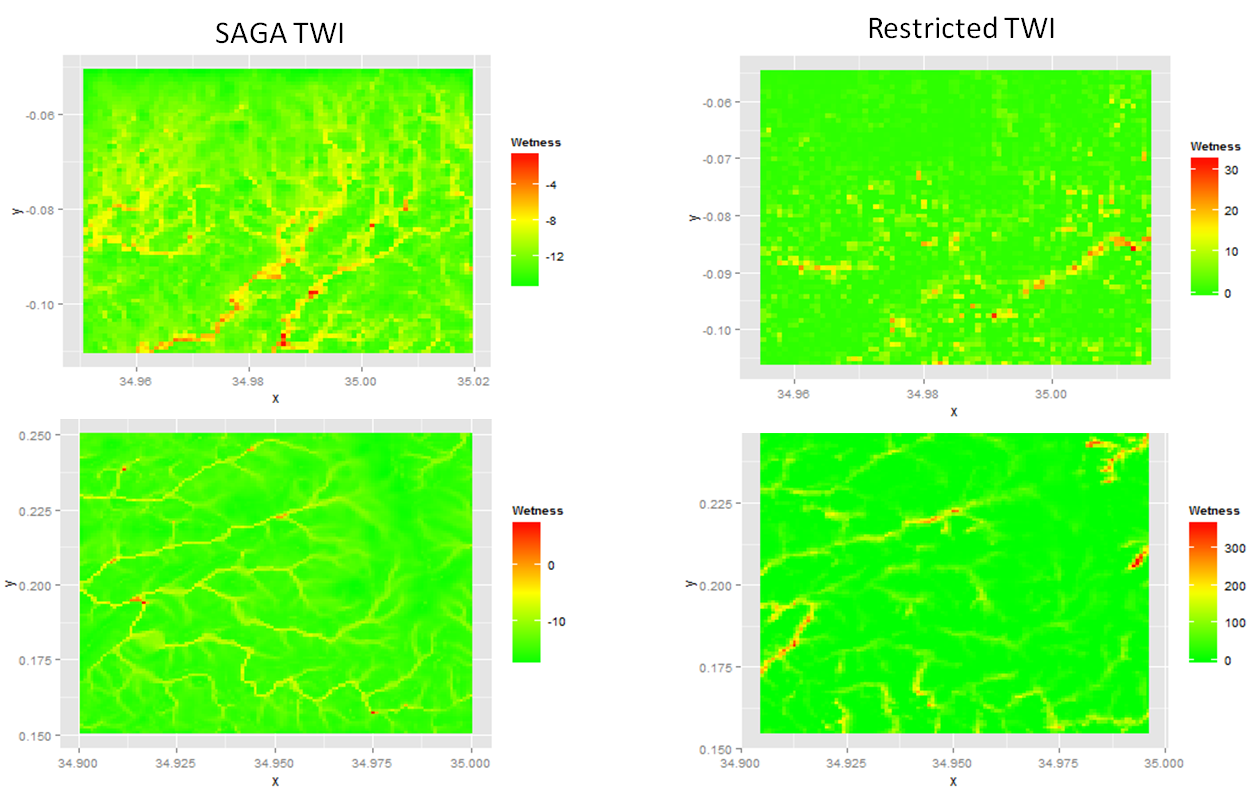


Figure 1. Comparison of the restricted TWI algorithm with the SAGA packaged algorithm.

The differential efficacy of treatment administration we found between the high and low sites is troubling. Blanket coverage of these treatments to every household in areas of high malaria transmission throughout the globe is extremely unlikely. Therefore, it is critically important that treatments are administered to households at greatest risk. Moreover, we feel that household risk should simultaneously account for both mosquito exposure and individual age-based risk so that those who are both likely to be bitten by a malaria harboring mosquito *and* are likely to suffer severe health outcomes from those bites are targeted for intervention. We believe a combination of current protocols and TWI based assessment of household locations can significantly improve administration of interventions.

# Literature Cited

Cohen, J.M., K.C. Ernst, K.A. Linblade, J.M. Vulule, C.C. John, and M. Wilson. 2008. Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western Kenyan highlands. Malaria Journal 7: 40.

Cohen, J.M., K.C. Ernst, K.A. Linblade, J.M. Vulule, C.C. John, and M. Wilson. 2010. Local topographic wetness indices predict household malaria risk better than land-use and land-cover in the western Kenya highlands. Malaria Journal 9: 328.

Gupta, S., R.W. Snow, C.A. Donnelly, K. Marsh, and C. Newbold. 1999. Immunity to non-cerebral severe malaria is acquired after one or two infections. Nature Medicine 5: 340-343.

Menendez, C., J. Ordi, M.R. Ismail, P.J. Ventura, J.J. Aponte, E Kahigwa, F. Font, and P.L. Alonso. 2000. The Impact of Placental Malaria on Gestational Age and Birth Weight. J Infectious Dis. 181 (5): 1740-1745.

Snow, R.W., M. Craig, U. Deichmann, and K. Marsh. 1999. Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population. Bull. World Health Organ. 77(8): 624-640.

# Appendix: R Code

**Session Information:**

R version 3.1.1 (2014-07-10)

Platform: x86\_64-w64-mingw32/x64 (64-bit)

locale:

[1] LC\_COLLATE=English\_United States.1252 LC\_CTYPE=English\_United States.1252

[3] LC\_MONETARY=English\_United States.1252 LC\_NUMERIC=C

[5] LC\_TIME=English\_United States.1252

attached base packages:

[1] splines grid stats graphics grDevices utils datasets methods base

other attached packages:

[1] SDMTools\_1.1-221 rgdal\_0.9-1 raster\_2.3-12 sp\_1.0-15 BDSS\_0.1.3

[6] pander\_0.3.8 rms\_4.2-0 SparseM\_1.05 Hmisc\_3.14-4 Formula\_1.1-2

[11] survival\_2.37-7 lattice\_0.20-29 rtf\_0.4-11 xtable\_1.7-3 MASS\_7.3-33

[16] ggmap\_2.3 ggplot2\_1.0.0 car\_2.0-20

loaded via a namespace (and not attached):

[1] cluster\_1.15.2 colorspace\_1.2-4 digest\_0.6.4 evaluate\_0.5.5

[5] formatR\_0.10 fortunes\_1.5-2 gtable\_0.1.2 htmltools\_0.2.4

[9] knitr\_1.6 labeling\_0.2 latticeExtra\_0.6-26 mapproj\_1.2-2

[13] maps\_2.3-7 munsell\_0.4.2 nnet\_7.3-8 plyr\_1.8.1

[17] png\_0.1-7 proto\_0.3-10 R.methodsS3\_1.6.1 R.oo\_1.18.0

[21] R.utils\_1.34.0 RColorBrewer\_1.0-5 Rcpp\_0.11.2 reshape2\_1.4

[25] RgoogleMaps\_1.2.0.6 rjson\_0.2.14 RJSONIO\_1.2-0.2 rmarkdown\_0.2.50

[29] scales\_0.2.4 stringr\_0.6.2 tools\_3.1.1 yaml\_2.1.13

**Code:**

## ----ReadIn,echo=FALSE,message=FALSE,warning=FALSE-----------------------

#read in and check data

rm(list=ls())

library(car)

library(ggplot2)

library(ggmap)

library(MASS)

longdat<-read.csv("C:/Classes/AppliedBiostat/Project/Domconv\_918.csv")[1:17823,]

longdat$site<-Recode(longdat$site,"c('High','HIGH')='High'")

longdat$Villcodef<-as.factor(longdat$Villcode)

longdat$Housef<-as.factor(longdat$House)

#create a variable which combines subcode, villcode, and house

longdat$household<-factor(with(longdat, paste(Subcode, Villcode, House, sep="\_")))

longdat$latlon<-factor(with(longdat, paste(Lat, Long, sep="\_")))

# There are more unique housholds than latlons!

# strsplit(unique(dat$latlon),split="\_") #get unique lat lons in a list

longdat$Age<-as.numeric(as.character(longdat$Age))#introducing NAs on purpose

longdat$HHhead<-Recode(longdat$HHhead,"c('F','M','N/A')='N'")

longdat$Relate<-as.numeric(as.character(Recode(longdat$Relate,"'N/A'=NA")))

longdat$HHspray<-Recode(longdat$HHspray,"c('','4','N/A')=NA")

longdat$NetHH<-Recode(longdat$NetHH,"'N/A'=NA")

longdat$NetHH<-droplevels(longdat$NetHH)

## ----eliminateDuplicates,echo=FALSE,eval=FALSE---------------------------

## #find duplicates entries within households

houses<-factor(unique(longdat$household))

N<-length(houses)

#loop through households and find entires that have unique age sex combos

longdat$age\_sex<-factor(paste(longdat$Age,longdat$Sex,sep="\_"))

subdat<-list(length=N)

for( i in 1:N){

#limit to 1 household

housedat<-longdat[longdat$household==houses[i],]

include<-which(!duplicated(housedat$age\_sex))

subdat[[i]]<-housedat[include,]

}

datunique<-do.call(rbind,subdat)

rm(subdat)

## ----summarizeHouseHolds, echo=FALSE, eval=FALSE-------------------------

## #This chunk is to create a new data set which is based on the household

numocc<-numFem<-numMale<-oldest<-youngest<-headsex<-headage<-net<-under5<-under1<-under18<-over65<-periph<-lat<-lon<-hspray<-site<-elev<-lang<-rep(0,length=N)

## #loop through each unique household and calculate summary variables

for (i in 1:N){

numocc[i]<-length(which(datunique$household==houses[i]))

numFem[i]<-sum(datunique$household==houses[i] & datunique$Sex=="F",na.rm=T)

numMale[i]<-sum(datunique$household==houses[i] & datunique$Sex=="M",na.rm=T)

oldest[i]<-max(datunique[datunique$household==houses[i],]$Age)

youngest[i]<-min(datunique[datunique$household==houses[i],]$Age)

if(length(datunique[(datunique$household==houses[i] & datunique$HHhead=="Y"),]$Sex)==1){

headsex[i]<-as.character(datunique[(datunique$household==houses[i] & datunique$HHhead=="Y"),]$Sex)

}else{headsex[i]<-NA}

if(length(datunique[datunique$household==houses[i]& datunique$HHhead=="Y",]$Age)==1){

headage[i]<-datunique[datunique$household==houses[i]& datunique$HHhead=="Y",]$Age

}else{headage[i]<-NA}

netans<-as.character(datunique[datunique$household==houses[i],]$NetHH)

net[i]<-sum(netans=="Y",na.rm=T)

under5[i]<-sum(datunique[datunique$household==houses[i],]$Age>1 & datunique[datunique$household==houses[i],]$Age<=5,na.rm=T)#children >1 and less than 5

under1[i]<-sum(datunique[datunique$household==houses[i],]$Age<=1,na.rm=T)

under18[i]<-sum(datunique[datunique$household==houses[i],]$Age<18,na.rm=T)

over65[i]<-sum(datunique[datunique$household==houses[i],]$Age>=65,na.rm=T)

family<-datunique[datunique$household==houses[i],]$Relate

periph[i]<-any(family>9)#classify any one more distant than grandchild as periphery

lat[i]<-datunique[datunique$household==houses[i],]$Lat[1]

lon[i]<-datunique[datunique$household==houses[i],]$Long[1]

spray<-datunique[datunique$household==houses[i],]$HHspray

hspray[i]<-sum(spray=="Y",na.rm=T)

site[i]<-datunique[datunique$household==houses[i],]$site[1]

elev[i]<-datunique[datunique$household==houses[i],]$Elevation[1]

lang[i]<-datunique[datunique$household==houses[i],]$Lang[1]

}

#create data frame from new data vectors

hdat<-data.frame(houses,numocc,numFem,numMale,oldest,youngest,headsex,headage,net,under5,under1,under18,over65,periph,lat,lon,hspray,site,elev,lang)

#save new data to csv file so I don't have to run this costly loop again

write.csv(hdat,"C:/Classes/AppliedBiostat/Project/housedata.csv")

## ----importNewData,echo=FALSE,message=FALSE,results='asis'---------------

rm(list=ls())

library(BDSS)

library(pander)

hdat<-read.csv("C:/Classes/AppliedBiostat/Project/housedata.csv")

sumdat<-hdat

sumdat$net<-factor(ifelse(is.na(sumdat$net),NA,ifelse(sumdat$net==0,"No","Yes")))

sumdat$under5f<-factor(ifelse(is.na(sumdat$under5),NA,ifelse(sumdat$under5>=1,"Yes","No")))

sumdat$under1f<-factor(ifelse(is.na(sumdat$under1),NA,ifelse(sumdat$under1>=1,"Yes","No")))

sumdat$under18f<-factor(ifelse(is.na(sumdat$under18),NA,ifelse(sumdat$under18>=1,"Yes","No")))

sumdat$over65f<-factor(ifelse(is.na(sumdat$under18),NA,ifelse(sumdat$over65>=1,"Yes","No")))

sumdat$periph<-factor(ifelse(is.na(sumdat$periph),NA,ifelse(sumdat$periph==1,"Yes","No")))

sumdat$hspray<-factor(ifelse(is.na(sumdat$hspray),NA,ifelse(sumdat$hspray==0,"No","Yes")))

sumdat$site<-factor(ifelse(is.na(sumdat$site),NA,ifelse(sumdat$site==1,"High","Low")))

sumdat[which(sumdat$site=="Low" & sumdat$lat>0),]$site<-"High"#correct mis-labelled site

## ----HouseSummary,echo=FALSE,results='asis',eval=FALSE-------------------

SummaryTable(data=sumdat,rowvars=c("under5f","under1f","under18f","over65f","periph","hspray","site","headsex"),row.names=c("Under 5","Under 1","Under 18","Peripheral Family","House Sprayed","Site","House Head Sex"),colvar="net",cont.vars=c("numocc","numFem","numMale","oldest","youngest","headage","elev"),output="rmarkdown")

## ----CalcAgeRisk,echo=FALSE----------------------------------------------

sumdat$arisk<-with(sumdat,(2\*under1)+under5+over65)

sumdat$ariskstd<-scale(log(sumdat$arisk+1),center=F)

#for sensitivity analysis exclude elderly group

sumdat$arisks<-with(sumdat,(2\*under1)+under5)

sumdat$ariskstds<-scale(log(sumdat$arisks+1),center=F)

## ----getDEM,echo=FALSE,message=FALSE,message=FALSE,results='hide'--------

library(raster,quietly=TRUE,warn.conflicts=FALSE)

library(ggplot2)

demlow<-getData(name='SRTM',download=FALSE,path="C:/Classes/AppliedBiostat/Project",lon=34.98,lat=-.08)

demlow2<-getData(name='SRTM',download=FALSE,path="C:/Classes/AppliedBiostat/Project",lon=35.98,lat=-.08)

#plot(dem)

demhigh<-getData(name='SRTM',download=FALSE,path="C:/Classes/AppliedBiostat/Project",lon=34.965,lat=.2)

lowel<-readGDAL("C:/Classes/AppliedBiostat/Project/srtm\_43\_13.tif")

lowel2<-readGDAL("C:/Classes/AppliedBiostat/Project/srtm\_44\_13.tif")

highel<-readGDAL("C:/Classes/AppliedBiostat/Project/srtm\_43\_12.tif")

#shrink area

lowext<-extent(34.950000,35.00042,-0.110000,-0.050000)

lowel<-crop(demlow,lowext,"C:/Classes/AppliedBiostat/Project/LowSiteData",overwrite=TRUE)

lowext2<-extent(35.000000,35.020000,-0.110000,-0.050000)

lowel2<-crop(demlow2,lowext2,"C:/Classes/AppliedBiostat/Project/LowSiteData2",overwrite=TRUE)

# image(lowel)

highext<-extent(34.900000,35.000417,0.150000,0.250000)

highel<-crop(demhigh,highext,"C:/Classes/AppliedBiostat/Project/HighSiteData",overwrite=TRUE)

lowelcomp<-merge(lowel,lowel2)

writeGDAL(as(lowelcomp,"SpatialPixelsDataFrame"),fname="LowDEM.tif")

writeGDAL(as(highel,"SpatialPixelsDataFrame"),fname="HighDEM.tif")

lowsite<-sumdat[sumdat$site=="Low",]

lowp<-data.frame(rasterToPoints(lowelcomp))

ggplot(data=lowp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1200, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=hspray),data=lowsite[-565,],alpha=0.4)+scale\_colour\_manual("Spray",values=c("red","blue"))

highsite<-sumdat[sumdat$site=="High",]

highp<-data.frame(rasterToPoints(highel))

ggplot(data=highp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1714, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=hspray),data=highsite,alpha=0.4)+scale\_colour\_manual("Spray",values=c("red","blue"))

## ----computeSWI,echo=FALSE,message=FALSE,warning=FALSE,results='hide'----

library(SDMTools)

#calculate elevation variance

#first get diff from mean to deterime if it is a dip or a peak

#create 3 different windows for calculating dips, this will account for drainage area to some extent

meanmat3<-matrix(1/9,3,3)

meanmat7<-matrix(1/49,7,7)

meanmat11<-matrix(1/121,11,11)

#find dips at these three different windows

meanel3<-focal(lowelcomp,w=meanmat3)#,pad=T,pad.value=mean(values(lowelcomp)))

meandiff3<-lowelcomp-meanel3

values(meandiff3)<-ifelse(values(meandiff3)<0,values(meandiff3),0)

meanel7<-focal(lowelcomp,w=meanmat7)#,pad=T,pad.value=mean(values(lowelcomp)))

meandiff7<-lowelcomp-meanel7

values(meandiff7)<-ifelse(values(meandiff7)<0,values(meandiff7),0)

meanel11<-focal(lowelcomp,w=meanmat11)#,pad=T,pad.value=mean(values(lowelcomp)))

meandiff11<-lowelcomp-meanel11

values(meandiff11)<-ifelse(values(meandiff11)<0,values(meandiff11),0)

lowdips<-meandiff3+meandiff7+meandiff11

#calculate aspect of each cell

lowasp<-aspect(lowelcomp,latlon=TRUE)

lowaspc<-lowasp

#categorize into quadrants

values(lowaspc)<-ifelse(values(lowasp)>45 & values(lowasp) <= 135,1,

ifelse(values(lowasp)>135 & values(lowasp) <= 225,2,

ifelse(values(lowasp)>225 & values(lowasp) <= 315,3,

ifelse(values(lowasp)>315 & values(lowasp) <=360,4,

ifelse(values(lowasp)>0 & values(lowasp) <= 45,4,NA)))))

#Calculate aspect variance

aspmat3<-matrix(1,3,3)

aspmat7<-matrix(1,7,7)

aspmat11<-matrix(1,11,11)

aspvar3<-focal(lowaspc,aspmat3,fun=var,na.rm=T)

aspvar7<-focal(lowaspc,aspmat7,fun=var,na.rm=T)

aspvar11<-focal(lowaspc,aspmat11,fun=var,na.rm=T)

lowaspvar<-aspvar3+aspvar7+aspvar11

lowtwi<-lowaspvar\*lowdips\*-1

#repeat for High site

meanel3<-focal(highel,w=meanmat3)#,pad=T,pad.value=mean(values(highel)))

meandiff3<-highel-meanel3

values(meandiff3)<-ifelse(values(meandiff3)<0,values(meandiff3),0)

meanel7<-focal(highel,w=meanmat7)#,pad=T,pad.value=mean(values(highel)))

meandiff7<-highel-meanel7

values(meandiff7)<-ifelse(values(meandiff7)<0,values(meandiff7),0)

meanel11<-focal(highel,w=meanmat11)#,pad=T,pad.value=mean(values(highel)))

meandiff11<-highel-meanel11

values(meandiff11)<-ifelse(values(meandiff11)<0,values(meandiff11),0)

highdips<-meandiff3+meandiff7+meandiff11

#Repeat for High site

highasp<-aspect(highel,latlon=TRUE)

highaspc<-highasp

#categorize into quadrants

values(highaspc)<-ifelse(values(highasp)>45 & values(highasp) <= 135,1,

ifelse(values(highasp)>135 & values(highasp) <= 225,2,

ifelse(values(highasp)>225 & values(highasp) <= 315,3,

ifelse(values(highasp)>315 & values(highasp) <=360,4,

ifelse(values(highasp)>0 & values(highasp) <= 45,4,NA)))))

#Calculate aspect variance

aspmat3<-matrix(1,3,3)

aspmat7<-matrix(1,7,7)

aspmat11<-matrix(1,11,11)

aspvar3<-focal(highaspc,aspmat3,fun=var,na.rm=T)

aspvar7<-focal(highaspc,aspmat7,fun=var,na.rm=T)

aspvar11<-focal(highaspc,aspmat11,fun=var,na.rm=T)

highaspvar<-aspvar3+aspvar7+aspvar11

hightwi<-highaspvar\*highdips\*-1

#plot restricted TWI

htwi<-as.data.frame(hightwi,xy=TRUE)

ltwi<-as.data.frame(lowtwi,xy=TRUE)

ggplot(data=ltwi[complete.cases(ltwi),],aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Wetness",low = "green", mid = "yellow", high = "red", midpoint =15, space = "rgb", na.value = "grey50", guide = "colourbar")

#

ggplot(data=htwi[complete.cases(htwi),],aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Wetness",low = "green", mid = "yellow", high = "red", midpoint =max(htwi$layer,na.rm=TRUE)/2, space = "rgb", na.value = "grey50", guide = "colourbar")

lowTWI<-readGDAL("C:/Classes/AppliedBiostat/Project/LowTWI.tif")#TWI calculated by SAGA

highTWI<-readGDAL("C:/Classes/AppliedBiostat/Project/HighTWI.tif")

lowt<-data.frame(lowTWI)

ggplot(data=lowt,aes(x=x,y=y))+geom\_raster(aes(fill=band1))+scale\_fill\_gradient2("Wetness",low = "green", mid = "yellow", high = "red", midpoint =-8, space = "rgb", na.value = "grey50", guide = "colourbar")#+ geom\_point(aes(x=lon, y=lat,color=hspray),data=lowsite[-565,])+scale\_colour\_manual("Spray",values=c("green","blue"))

hight<-data.frame(highTWI)

ggplot(data=hight,aes(x=x,y=y))+geom\_raster(aes(fill=band1))+scale\_fill\_gradient2("Wetness",low = "green", mid = "yellow", high = "red", midpoint =-5, space = "rgb", na.value = "grey50", guide = "colourbar")#+ geom\_point(aes(x=lon, y=lat,color=hspray),data=highsite)+scale\_colour\_manual("Spray",values=c("red","blue"))

## ----smoothTWI,echo=FALSE------------------------------------------------

#use focal to create moving window with gaussian weights for assiging household risk

#create filter for focal window

r<-raster(ncols=36,nrows=36,xmn=0)

#convert TWI to raster

lowtr<-raster(lowTWI,layer="band1")

hightr<-raster(highTWI,layer="band1")

fmat<-focalWeight(r,d=10,type='Gauss')

mriskLow<-focal(lowtr,fmat,filename='C:/Classes/AppliedBiostat/Project/smoothLowTWI',pad=T,padValue=median(lowt$band1),overwrite=T)

mriskHigh<-focal(hightr,fmat,filename='C:/Classes/AppliedBiostat/Project/smoothHighTWI',pad=T,padValue=median(hight$band1),overwrite=T)

lowtrp<-as.data.frame(mriskLow,xy=T)

hightrp<-as.data.frame(mriskHigh,xy=T)

ggplot(data=lowtrp[complete.cases(lowtrp),],aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Mosquito Risk",low = "green", mid = "yellow", high = "red", midpoint = -12, space = "rgb", na.value = "grey50", guide = "colourbar")+theme(legend.position='bottom')+ geom\_point(aes(x=lon, y=lat),data=lowsite[-565,],alpha=0.9)

ggplot(data=hightrp[complete.cases(hightrp),],aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Mosquito Risk",low = "green", mid = "yellow", high = "red", midpoint = median(hightrp$layer,na.rm=T), space = "rgb", na.value = "grey50", guide = "colourbar")+theme(legend.position='bottom')+ geom\_point(aes(x=lon, y=lat),data=highsite,alpha=0.8)

## ----assignRisktoHouse,echo=FALSE----------------------------------------

lowsite$mrisk<-extract(mriskLow,y=cbind(lowsite$lon,lowsite$lat))

highsite$mrisk<-extract(mriskHigh,y=cbind(highsite$lon,highsite$lat))

#Standardize the risk scores

lowsite$mriskpos<-lowsite$mrisk+abs(min(lowsite$mrisk,na.rm=T))

highsite$mriskpos<-highsite$mrisk+abs(min(highsite$mrisk,na.rm=T))

lowsite$mriskstd<-scale(log(lowsite$mriskpos+1),center=F)

highsite$mriskstd<-scale(log(highsite$mriskpos+1),center=F)

## ----plotMrisk,echo=FALSE,eval=FALSE-------------------------------------

ggplot(data=lowp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1200, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=mriskstd),data=lowsite[-565,],alpha=0.9)+scale\_colour\_gradient("Mosquito Risk",low="blue",high="red")

ggplot(data=highp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1714, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=mriskstd),data=highsite,alpha=0.8)+scale\_colour\_gradient("Mosquito Risk",low="blue",high="red")

## ----plotArisk,echo=FALSE,eval=FALSE-------------------------------------

ggplot(data=lowp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1200, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=arisk),data=lowsite[-565,],alpha=0.9)+scale\_colour\_gradient("Age Risk",low="blue",high="red")

ggplot(data=highp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1714, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=arisk),data=highsite,alpha=0.8)+scale\_colour\_gradient("Age Risk",low="blue",high="red")

## ----CombinedRisk,echo=FALSE---------------------------------------------

lowsite$comrsk<-lowsite$mriskstd+lowsite$arisk

highsite$comrsk<-highsite$mriskstd+highsite$arisk

## ----plotCombrisk,echo=FALSE,fig.height=3,fig.width=3,eval=FALSE---------

ggplot(data=lowp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1200, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=comrsk),data=lowsite[-565,],alpha=0.9)+scale\_colour\_gradient("Combined Risk",low="blue",high="red")+theme(legend.position='bottom',legend.box='horizontal')

ggplot(data=highp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1714, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=comrsk),data=highsite,alpha=0.8)+scale\_colour\_gradient("Combined Risk",low="blue",high="red")+theme(legend.position='bottom',legend.box='horizontal')

## ----checkLinearityAssumption,echo=FALSE,results='hide',eval=FALSE-------

lowsite$hspray<-ifelse(lowsite$hspray=="No",0,1)

lowsite$net<-ifelse(lowsite$net=="No",0,1)

highsite$hspray<-ifelse(highsite$hspray=="No",0,1)

highsite$net<-ifelse(highsite$net=="No",0,1)

rcspline.plot(lowsite$comrsk,lowsite$net,model="logistic",xlab="Combined Risk",ylab="Log Odds Net Use")#evidence of non-linearity (u-shape)

rcspline.plot(highsite$comrsk,highsite$net,model="logistic",xlab="Combined Risk")

rcspline.plot(lowsite$comrsk,lowsite$hspray,model="logistic")

rcspline.plot(highsite$comrsk,highsite$hspray,model="logistic")

rcspline.plot(lowsite$arisk,lowsite$net,model="logistic")

rcspline.plot(highsite$arisk,highsite$net,model="logistic")

rcspline.plot(lowsite$arisk,lowsite$hspray,model="logistic")

rcspline.plot(highsite$arisk,highsite$hspray,model="logistic")

rcspline.plot(lowsite$mrisk,lowsite$net,model="logistic")

rcspline.plot(highsite$mrisk,highsite$net,model="logistic",xlab="Mosquito Risk")#evidence of non-linearity increasing, leveling off then slight decrease

rcspline.plot(lowsite$mrisk,lowsite$hspray,model="logistic")

rcspline.plot(highsite$mrisk,highsite$hspray,model="logistic",xlab="Mosquito Risk")#evidence of non-linearity (increasing then leveling off)

## ----model1,echo=FALSE---------------------------------------------------

m1ls<-glm(hspray~comrsk,data=lowsite,family="binomial")

m1ls.est<-c(exp(summary(m1ls)$coef[2,1]),exp(summary(m1ls)$coef[2,1]-1.96\*summary(m1ls)$coef[2,2]),exp(summary(m1ls)$coef[2,1]+1.96\*summary(m1ls)$coef[2,2]))

m1ln<-glm(net~comrsk,data=lowsite,family="binomial")

m1ln.est<-c(exp(summary(m1ln)$coef[2,1]),exp(summary(m1ln)$coef[2,1]-1.96\*summary(m1ln)$coef[2,2]),exp(summary(m1ln)$coef[2,1]+1.96\*summary(m1ln)$coef[2,2]))

#adjust for non-linearity

lq1<-quantile(lowsite$comrsk,prob=.25)

lq2<-quantile(lowsite$comrsk,prob=.5)

lq3<-quantile(lowsite$comrsk,prob=.75)

lowsite$comrsk\_q<-ifelse(lowsite$comrsk <= lq1,1,ifelse(lowsite$comrsk>lq1 & lowsite$comrsk<=lq2,2,ifelse(lowsite$comrsk>lq2 & lowsite$comrsk<=lq3,3,4)))

m1lnq<-glm(net~factor(comrsk\_q),data=lowsite,family="binomial")

# summary(m1lnq)

m1hs<-glm(hspray~comrsk,data=highsite,family="binomial")

m1hs.est<-c(exp(summary(m1hs)$coef[2,1]),exp(summary(m1hs)$coef[2,1]-1.96\*summary(m1hs)$coef[2,2]),exp(summary(m1hs)$coef[2,1]+1.96\*summary(m1hs)$coef[2,2]))

m1hn<-glm(net~comrsk,data=highsite,family="binomial")

m1hn.est<-c(exp(summary(m1hn)$coef[2,1]),exp(summary(m1hn)$coef[2,1]-1.96\*summary(m1hn)$coef[2,2]),exp(summary(m1hn)$coef[2,1]+1.96\*summary(m1hn)$coef[2,2]))

Treatment<-rep(c("Net","Spray"),2)

Site<-c("High","","Low","")

outmat<-round(matrix(c(m1hn.est,m1hs.est,m1ln.est,m1ls.est),4,3,byrow=T),2)

colnames(outmat)<-c("OR","Lower 95% CI","Upper 95% CI")

obj1out<-cbind(Site,Treatment,outmat)

## ----Obj1Table,echo=FALSE,results='asis'---------------------------------

pandoc.table(obj1out,type="rmarkdown",split.tables=Inf)

## ----SepModels,echo=FALSE------------------------------------------------

#Separate model for each risk type

m2als<-glm(hspray~arisk,data=lowsite,family="binomial")

m2als.est<-c(exp(summary(m2als)$coef[2,1]),exp(summary(m2als)$coef[2,1]-1.96\*summary(m2als)$coef[2,2]),exp(summary(m2als)$coef[2,1]+1.96\*summary(m2als)$coef[2,2]))

m2aln<-glm(net~arisk,data=lowsite,family="binomial")

m2aln.est<-c(exp(summary(m2aln)$coef[2,1]),exp(summary(m2aln)$coef[2,1]-1.96\*summary(m2aln)$coef[2,2]),exp(summary(m2aln)$coef[2,1]+1.96\*summary(m2aln)$coef[2,2]))

m2mls<-glm(hspray~mrisk,data=lowsite,family="binomial")

m2mls.est<-c(exp(summary(m2mls)$coef[2,1]),exp(summary(m2mls)$coef[2,1]-1.96\*summary(m2mls)$coef[2,2]),exp(summary(m2mls)$coef[2,1]+1.96\*summary(m2mls)$coef[2,2]))

m2mln<-glm(net~mrisk,data=lowsite,family="binomial")

m2mln.est<-c(exp(summary(m2mln)$coef[2,1]),exp(summary(m2mln)$coef[2,1]-1.96\*summary(m2mln)$coef[2,2]),exp(summary(m2mln)$coef[2,1]+1.96\*summary(m2mln)$coef[2,2]))

m2ahs<-glm(hspray~arisk,data=highsite,family="binomial")

m2ahs.est<-c(exp(summary(m2ahs)$coef[2,1]),exp(summary(m2ahs)$coef[2,1]-1.96\*summary(m2ahs)$coef[2,2]),exp(summary(m2ahs)$coef[2,1]+1.96\*summary(m2ahs)$coef[2,2]))

m2ahn<-glm(net~arisk,data=highsite,family="binomial")

m2ahn.est<-c(exp(summary(m2ahn)$coef[2,1]),exp(summary(m2ahn)$coef[2,1]-1.96\*summary(m2ahn)$coef[2,2]),exp(summary(m2ahn)$coef[2,1]+1.96\*summary(m2ahn)$coef[2,2]))

m2mhs<-glm(hspray~mrisk,data=highsite,family="binomial")

m2mhs.est<-c(exp(summary(m2mhs)$coef[2,1]),exp(summary(m2mhs)$coef[2,1]-1.96\*summary(m2mhs)$coef[2,2]),exp(summary(m2mhs)$coef[2,1]+1.96\*summary(m2mhs)$coef[2,2]))

m2mhn<-glm(net~mrisk,data=highsite,family="binomial")

m2mhn.est<-c(exp(summary(m2mhn)$coef[2,1]),exp(summary(m2mhn)$coef[2,1]-1.96\*summary(m2mhn)$coef[2,2]),exp(summary(m2mhn)$coef[2,1]+1.96\*summary(m2mhn)$coef[2,2]))

sepmat<-round(matrix(c(m2ahn.est,m2mhn.est,m2ahs.est,m2mhs.est,m2aln.est,m2mln.est,m2als.est,m2mls.est),4,6,byrow=T),2)

colnames(sepmat)<-c(rep("Age Risk",3),rep("Mosquito Risk",3))

est<-c("","",rep(c("OR","Lower 95% CI","Upper 95% CI"),2))

sepout<-cbind(Site,Treatment,sepmat)

sepout<-rbind(est,sepout)

rownames(sepout)<-1:nrow(sepout)

## ----objective2Table,echo=FALSE,results='asis'---------------------------

pandoc.table(sepout,type="rmarkdown",split.tables=Inf)

## ----smoothTWI\_sens,echo=FALSE-------------------------------------------

#use focal to create moving window with gaussian weights for assiging household risk

#create filter for focal window

r<-raster(ncols=36,nrows=36,xmn=0)

#convert TWI to raster

lowtr\_s<-lowtwi

hightr\_s<-hightwi

fmat<-focalWeight(r,d=10,type='Gauss')

mriskLows<-focal(lowtr\_s,fmat,filename='smoothLowTWI\_sens',pad=T,padValue=median(lowt$band1),overwrite=T)

mriskHighs<-focal(hightr\_s,fmat,filename='smoothHighTWI\_sens',pad=T,padValue=median(hight$band1),overwrite=T)

lowtrps<-as.data.frame(mriskLows,xy=T)

hightrps<-as.data.frame(mriskHighs,xy=T)

ggplot(data=lowtrps[complete.cases(lowtrps),],aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Mosquito Risk",low = "green", mid = "yellow", high = "red", midpoint = median(lowtrps$layer,na.rm=T), space = "rgb", na.value = "grey50", guide = "colourbar")+theme(legend.position='bottom')+ geom\_point(aes(x=lon, y=lat),data=lowsite[-565,],alpha=0.9)

ggplot(data=hightrps[complete.cases(hightrps),],aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Mosquito Risk",low = "green", mid = "yellow", high = "red", midpoint = median(hightrps$layer,na.rm=T), space = "rgb", na.value = "grey50", guide = "colourbar")+theme(legend.position='bottom')+ geom\_point(aes(x=lon, y=lat),data=highsite[-2537,],alpha=0.8)

## ----assignRisktoHouse\_sens,echo=FALSE-----------------------------------

lowsite$mrisks<-extract(mriskLows,y=cbind(lowsite$lon,lowsite$lat))

highsite$mrisks<-extract(mriskHighs,y=cbind(highsite$lon,highsite$lat))

#Standardize the risk scores

lowsite$mriskstds<-scale(log(lowsite$mrisks+1),center=F)

highsite$mriskstds<-scale(log(highsite$mrisks+1),center=F)

## ----plotMrisk\_sens,echo=FALSE,eval=FALSE--------------------------------

ggplot(data=lowp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1200, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=mrisks),data=lowsite,alpha=0.9)+scale\_colour\_gradient("Mosquito Risk",low="blue",high="red")

ggplot(data=highp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1714, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=mrisks),data=highsite,alpha=0.8)+scale\_colour\_gradient("Mosquito Risk",low="blue",high="red")

## ----CombinedRisk\_sens,echo=FALSE----------------------------------------

lowsite$comrsks<-scale(log(lowsite$mrisks+1),center=F)+scale(log(lowsite$arisk+1),center=F)

highsite$comrsks<-scale(log(highsite$mrisks+1),center=F)+scale(log(highsite$arisk+1),center=F)

## ----plotCombrisk\_sens,echo=FALSE,fig.height=7,fig.width=7,eval=FALSE----

ggplot(data=lowp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1200, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=comrsks),data=lowsite,alpha=0.9)+scale\_colour\_gradient("Combined Risk",low="blue",high="red")+theme(legend.position='bottom',legend.box='horizontal')

ggplot(data=highp,aes(x=x,y=y))+geom\_raster(aes(fill=layer))+scale\_fill\_gradient2("Elevation",low = "brown", mid = "yellow", high = "green", midpoint = 1714, space = "rgb", na.value = "grey50", guide = "colourbar")+ geom\_point(aes(x=lon, y=lat,color=comrsks),data=highsite,alpha=0.8)+scale\_colour\_gradient("Combined Risk",low="blue",high="red")+theme(legend.position='bottom',legend.box='horizontal')

## ----checkLinearityAssumption\_sens,echo=FALSE,results='hide',eval=FALSE----

rcspline.plot(lowsite$comrsks,lowsite$net,model="logistic",xlab="Combined Risk",ylab="Log Odds Net Use")#evidence of non-linearity (u-shape)

rcspline.plot(highsite$comrsks,highsite$net,model="logistic",xlab="Combined Risk")

rcspline.plot(lowsite$comrsks,lowsite$hspray,model="logistic")

rcspline.plot(highsite$comrsks,highsite$hspray,model="logistic")

rcspline.plot(lowsite$arisk,lowsite$net,model="logistic")

rcspline.plot(highsite$arisk,highsite$net,model="logistic")

rcspline.plot(lowsite$arisk,lowsite$hspray,model="logistic")

rcspline.plot(highsite$arisk,highsite$hspray,model="logistic")

rcspline.plot(lowsite$mrisks,lowsite$net,model="logistic")

rcspline.plot(highsite$mrisks,highsite$net,model="logistic",xlab="Mosquito Risk")

rcspline.plot(lowsite$mrisks,lowsite$hspray,model="logistic")

rcspline.plot(highsite$mrisks,highsite$hspray,model="logistic",xlab="Mosquito Risk")#evidence of non-linearity (increasing then leveling off)

## ----model1\_sens,echo=FALSE----------------------------------------------

m1ls<-glm(hspray~comrsks,data=lowsite,family="binomial")

m1ls.est<-c(exp(summary(m1ls)$coef[2,1]),exp(summary(m1ls)$coef[2,1]-1.96\*summary(m1ls)$coef[2,2]),exp(summary(m1ls)$coef[2,1]+1.96\*summary(m1ls)$coef[2,2]))

m1ln<-glm(net~comrsks,data=lowsite,family="binomial")

m1ln.est<-c(exp(summary(m1ln)$coef[2,1]),exp(summary(m1ln)$coef[2,1]-1.96\*summary(m1ln)$coef[2,2]),exp(summary(m1ln)$coef[2,1]+1.96\*summary(m1ln)$coef[2,2]))

m1hs<-glm(hspray~comrsks,data=highsite,family="binomial")

m1hs.est<-c(exp(summary(m1hs)$coef[2,1]),exp(summary(m1hs)$coef[2,1]-1.96\*summary(m1hs)$coef[2,2]),exp(summary(m1hs)$coef[2,1]+1.96\*summary(m1hs)$coef[2,2]))

m1hn<-glm(net~comrsks,data=highsite,family="binomial")

m1hn.est<-c(exp(summary(m1hn)$coef[2,1]),exp(summary(m1hn)$coef[2,1]-1.96\*summary(m1hn)$coef[2,2]),exp(summary(m1hn)$coef[2,1]+1.96\*summary(m1hn)$coef[2,2]))

Treatment<-rep(c("Net","Spray"),2)

Site<-c("High","","Low","")

outmats<-round(matrix(c(m1hn.est,m1hs.est,m1ln.est,m1ls.est),4,3,byrow=T),2)

colnames(outmats)<-c("OR","Lower 95% CI","Upper 95% CI")

obj1outs<-cbind(Site,Treatment,outmat,outmats)

## ----Obj1Table\_sens,echo=FALSE,results='asis'----------------------------

pandoc.table(obj1outs,type="rmarkdown",split.tables=Inf)

## ----SepModels\_sens,echo=FALSE-------------------------------------------

#Separate model for each risk type

m2als<-glm(hspray~arisk,data=lowsite,family="binomial")

m2als.est<-c(exp(summary(m2als)$coef[2,1]),exp(summary(m2als)$coef[2,1]-1.96\*summary(m2als)$coef[2,2]),exp(summary(m2als)$coef[2,1]+1.96\*summary(m2als)$coef[2,2]))

m2aln<-glm(net~arisk,data=lowsite,family="binomial")

m2aln.est<-c(exp(summary(m2aln)$coef[2,1]),exp(summary(m2aln)$coef[2,1]-1.96\*summary(m2aln)$coef[2,2]),exp(summary(m2aln)$coef[2,1]+1.96\*summary(m2aln)$coef[2,2]))

m2mls<-glm(hspray~arisks,data=lowsite,family="binomial")

m2mls.est<-c(exp(summary(m2mls)$coef[2,1]),exp(summary(m2mls)$coef[2,1]-1.96\*summary(m2mls)$coef[2,2]),exp(summary(m2mls)$coef[2,1]+1.96\*summary(m2mls)$coef[2,2]))

m2mln<-glm(net~arisks,data=lowsite,family="binomial")

m2mln.est<-c(exp(summary(m2mln)$coef[2,1]),exp(summary(m2mln)$coef[2,1]-1.96\*summary(m2mln)$coef[2,2]),exp(summary(m2mln)$coef[2,1]+1.96\*summary(m2mln)$coef[2,2]))

m2ahs<-glm(hspray~arisk,data=highsite,family="binomial")

m2ahs.est<-c(exp(summary(m2ahs)$coef[2,1]),exp(summary(m2ahs)$coef[2,1]-1.96\*summary(m2ahs)$coef[2,2]),exp(summary(m2ahs)$coef[2,1]+1.96\*summary(m2ahs)$coef[2,2]))

m2ahn<-glm(net~arisk,data=highsite,family="binomial")

m2ahn.est<-c(exp(summary(m2ahn)$coef[2,1]),exp(summary(m2ahn)$coef[2,1]-1.96\*summary(m2ahn)$coef[2,2]),exp(summary(m2ahn)$coef[2,1]+1.96\*summary(m2ahn)$coef[2,2]))

m2mhs<-glm(hspray~arisks,data=highsite,family="binomial")

m2mhs.est<-c(exp(summary(m2mhs)$coef[2,1]),exp(summary(m2mhs)$coef[2,1]-1.96\*summary(m2mhs)$coef[2,2]),exp(summary(m2mhs)$coef[2,1]+1.96\*summary(m2mhs)$coef[2,2]))

m2mhn<-glm(net~arisks,data=highsite,family="binomial")

m2mhn.est<-c(exp(summary(m2mhn)$coef[2,1]),exp(summary(m2mhn)$coef[2,1]-1.96\*summary(m2mhn)$coef[2,2]),exp(summary(m2mhn)$coef[2,1]+1.96\*summary(m2mhn)$coef[2,2]))

sepmats<-round(matrix(c(m2ahn.est,m2mhn.est,m2ahs.est,m2mhs.est,m2aln.est,m2mln.est,m2als.est,m2mls.est),4,6,byrow=T),2)

colnames(sepmats)<-c(rep("Age Risk (<5 or >65)",3),rep("Age Risk (<5)",3))

est<-c("","",rep(c("OR","Lower 95% CI","Upper 95% CI"),2))

sepouts<-cbind(Site,Treatment,sepmats)

sepouts<-rbind(est,sepouts)

rownames(sepouts)<-1:nrow(sepouts)

## ----objective2Table\_sens,echo=FALSE,results='asis'----------------------

pandoc.table(sepouts,type="rmarkdown",split.tables=Inf)