# Leptospirosis Risk Factors in Kenyan Pastoral Landscapes

Aditya Prabaswara Mardjikoen*

School of Mathematics, University of Edinburgh, Edinburgh, U.K.

* A.P.Mardjikoen@sms.ed.ac.uk

## 1. Introduction

A neglected but re-emerging zoonotic disease, leptospirosis is a global public health priority because of the potential impact on both human and animal health [1]. Recent decades have seen more outbreaks of these epidemics, particularly in developing countries [2]. According to numerous studies [1, 3, 4], leptospirosis is becoming more common due to factors such as outdoor work, animal contact, poor sanitation, and climate change. However, less work has been done to identify vulnerable individuals, especially in low-income rural areas in developing countries. Therefore, we aim to identify individual risk factors associated with leptospirosis infections in Kenyan rural areas. This study used mixed effects model [5, 6] and data provided by the International Livestock Research Institute (ILRI).

## 2. Fixed and Random Effects

The mixed effects model, or mixed model [5], is a regression models that contains both fixed effects and random effects. As [7] suggests, we define effects (or coefficients) in this model as constant (fixed effects) if they are similar for all groups in a population, and varying (random effects) if they are possible to vary between groups to groups.

## 3. Model

- **(Binary Logistic Model)** Define $Y_{ijk}$ as the ELISA (enzyme-linked immunosorbent assay) test results for person $i$ at village $j$ and household $k$ such that $Y_{ijk}$ is 1 if the test results is positive and 0 if otherwise. We denoted $p_{ijk}$ as the probability that $Y_{ijk}$ is 1.
- **(Fixed Effects Terms)** Define $X_{sijk}$ as the fixed effects variables where $s = 1, \ldots, n$. We denoted $\alpha$ as the fixed intercept and $\beta_s$ as the fixed slope for $X_{sijk}$.
- **(Random Effects Terms)** Let $\gamma_j$ and $\lambda_k$ be the random intercept for village and household respectively.
- **(Mixed Effects Model)** For inference, we follow the generalized linear mixed model (GLMM) [6] below to identify the risk factors of leptospirosis seropositivity.

**Binomial Generalized Linear Mixed Model**

$$\begin{aligned}
&\text{(Sampling Distribution)} \quad Y_{ijk} \sim \text{Bernoulli}(p_{ijk}),\\
&\text{(Model Specification)} \quad \text{logit}\,(p_{ijk}) = \alpha + \sum_{s=1}^{n} \beta_s X_{sijk} + \gamma_j + \lambda_k,\\
&\text{(Random Intercept)} \quad \gamma_j \sim N(0, \sigma_\gamma^2), \quad \lambda_k \sim N(0, \sigma_\lambda^2).
\end{aligned}$$

## 4. Data

The ILRI dataset was gathered in Tana River County, Kenya, in 2013 and 2014. It includes the following details:

- Individuals and household heads' age, gender, and behavioural data.
- Information about the individual's household location, jobs, and contact with animals.
- Results of leptospirosis ELISA tests from individuals living in a particular household.

Most individuals in the dataset are tested negative for leptospirosis using ELISA.

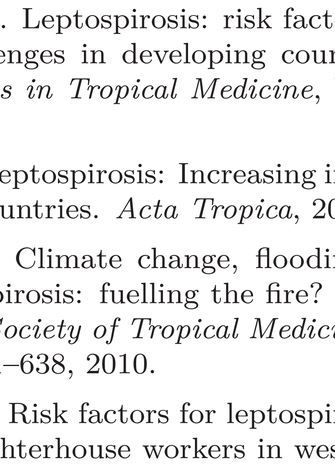


Fig 1. **Distribution of ELISA test results.**

We remove duplicated observations, observations with fewest household head jobs, and observations with missing data before fitting the model.

## 5. Results

The Odds Ratios (OR) [8] for an outcome $D$ associated with exposure $E$ is defined as follows:

$$\text{OR} = \frac{P(D|E)}{P(\text{not } D|E)} \times \frac{P(\text{not } D|\text{not } E)}{P(D|\text{not } E)},$$

where $P(\cdot)$ represents probability. The OR values can be interpreted as follows [8]:

- OR $= 1$ suggest independence of $D$ and $E$.
- OR $> 1$ suggest greater risk of $D$ if $E$ exists.
- OR $< 1$ suggest lower risk of $D$ if $E$ exists.

Our model explored the risk factors for leptospirosis seropositivity [4] based on the OR. We identified significant risk factors in our model based on their p-values $< 0.05$ and their 95% confidence intervals excluded 1. Our analysis suggests:

- Risk factors that were significant for leptospirosis seropositivity are females, increasing individual age, larger families, and living with a pastoralist household heads.
- Protective factors that were significant for leptospirosis seropositivity are males, living in a high-altitude village, and living with older household heads.
- Land use is not significantly associated with leptospirosis seropositivity.

| Variable | OR[1] | 95% CI[1] | p-value |
|---|---|---|---|
| Individual age | 1.006 | 1.003-1.01 | < 0.001 |
| Village altitude | 0.989 | 0.986-0.992 | < 0.001 |
| Individual gender | | | |
| Female[RC] | – | – | |
| Male | 0.62 | 0.44-0.88 | 0.007 |
| Household head age | 0.991 | 0.988-0.995 | < 0.001 |
| Family size | 1.037 | 1.03-1.04 | < 0.001 |
| Household head occupation | | | |
| Farmer[RC] | – | – | |
| Pastoralist | 5.82 | 1.87-18.13 | 0.002 |
| Land use | | | |
| Irrigation[RC] | – | – | |
| Pastoral | 0.37 | 0.12-1.19 | 0.1 |
| Riverine | 0.48 | 0.17-1.38 | 0.2 |

[1] OR = Odds Ratios, CI = Confidence Interval
[RC] Reference Category

Tbl 1. **Odds Ratios from mixed model analysis.**

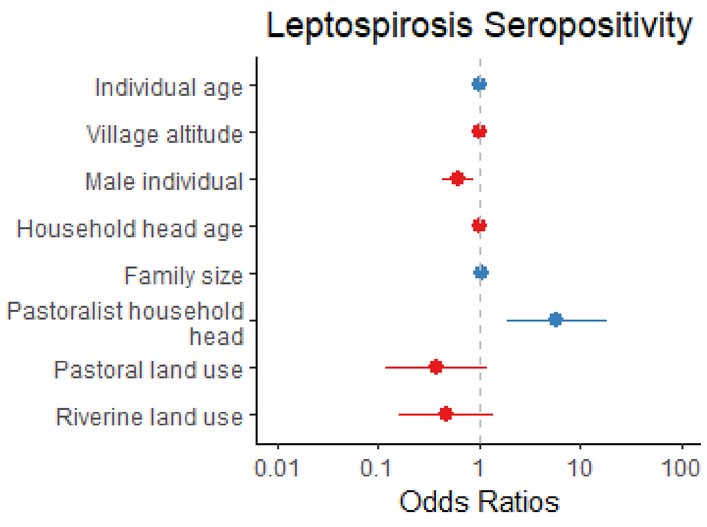


Fig 2. **Mixed model estimated Odds Ratios plot.** The estimated Odds Ratios (point) and its 95% confidence intervals (horizontal line).

## 6. References

[1] C. Goarant. Leptospirosis: risk factors and management challenges in developing countries. *Research and Reports in Tropical Medicine*, Volume 7:49–62, 2016.

[2] Soo *et al.* Leptospirosis: Increasing importance in developing countries. *Acta Tropica*, 201:105183, 2020.

[3] Lau *et al.* Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104(10):631–638, 2010.

[4] Cook *et al.* Risk factors for leptospirosis seropositivity in slaughterhouse workers in western kenya. *Occupational and Environmental Medicine*, 74(5):357–365, 2017.

[5] J.J. Faraway. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2016.

[6] Zuur *et al. Mixed Effects Models and Extensions in Ecology with R*. Statistics for Biology and Health. Springer, New York, 2009.

[7] A. Gelman. Analysis of variance—why it is more important than ever. *The annals of statistics*, 33(1):1–53, 2005.

[8] N.P. Jewell. *Statistics for Epidemiology*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, Boca Rotan, 2003.