

Supplemental Material 1: Statistical model

To understand the population and social determinants of morbidity, we fitted a mixed effects multivariate multinomial logistic model. This model allows us to investigate the relationship between multiple diagnosed conditions (malaria, gastric acid-related conditions, lower respiratory tract infections, urinary tract infections, coryza (common cold)) and various demographic and socioeconomic factors such as age groups, education level, wealth status, household size, number of visits to the facility and year of diagnosis, with coryza (common cold) as the reference category. This statistical approach allows for the simultaneous comparison of the odds of different diagnosed conditions relative to the reference category while accounting for multiple demographic and socioeconomic factors (predictor) ¹. The patient-specific, relating to a particular visit to the health facility, was included to account for the correlation due to multiple visits, possible multiple diagnosis.

Let Y_{ijk} and π_{ijk} be the diagnosed disease status and probability of being diagnosed with a particular disease, respectively, i.e., malaria ($k = 1$), lower respiratory tract infection ($k = 2$), gastric acid related ($k = 3$), urinary tract infection ($k = 4$), and coryza (common cold) ($k = 5$) of patient $i = 1, \dots$, in visit $j = 1, \dots$. We assume that Y_{ijk} follows a multinomial distribution, i.e., $Y_{ijk} \sim \text{MN}(1, \pi_{ijk})$, where $\pi_{ijk} = (\pi_{ij1}, \dots, \pi_{ij5})^T$ is the probability of patient i being diagnosed with condition $k \in 1, \dots, 5$ in visit j . The probability of being diagnosed with a particular disease, $k \in 1, \dots, 4$, relative to coryza (common cold) ($k = 5$), in visit j is defined as

$$\pi_{ijk} = \frac{\eta_{ijk}}{1 + \sum_{l=1}^k \exp(\eta_{ijl})}, \quad k = 1, 2, 3, 4$$

where η_{ijk} is the linear predictor given as

$$\eta_{ijk} = \beta_{0k} + \beta_{1k} \cdot \text{age group}_{ij} + \beta_{2k} \cdot \text{education level}_{ij} + \beta_{3k} \cdot \text{wealth status}_{ij} + \beta_{4k} \cdot \text{household size}_{ij} + \beta_{5k} \cdot \text{number of visits}_{ij} + \beta_{6k} \cdot \text{year of diagnosis}_{ij} + \tau_{ik}$$

and τ_{ik} is patient-specific random effects accounting for the correlation due to repeated measurements. For the reference category, $k = 5$, $\pi_{ij5} = 1 - \sum_{k=1}^4 \pi_{ijk}$.

More specifically, the linear predictors for each of the diagnosis category, other than Coryza, are specified as follows:

Malaria:

$$\eta_{ij,\text{Malaria}} = \beta_{0,\text{Malaria}} + \beta_{1,\text{Malaria}} \cdot \text{age group}_{ij} + \beta_{2,\text{Malaria}} \cdot \text{education level}_{ij} + \beta_{3,\text{Malaria}} \cdot \text{wealth status}_{ij} + \beta_{4,\text{Malaria}} \cdot \text{household size}_{ij} + \beta_{5,\text{Malaria}} \cdot \text{number of visits}_{ij} + \beta_{6,\text{Malaria}} \cdot \text{year of diagnosis}_{ij} + \tau_{i,\text{Malaria}}$$

Lower Respiratory Tract Infection (LRTI):

$$\eta_{ij,\text{LRTI}} = \beta_{0,\text{LRTI}} + \beta_{1,\text{LRTI}} \cdot \text{age group}_{ij} + \beta_{2,\text{LRTI}} \cdot \text{education level}_{ij} + \beta_{3,\text{LRTI}} \cdot \text{wealth status}_{ij} + \beta_{4,\text{LRTI}} \cdot \text{household size}_{ij} + \beta_{5,\text{LRTI}} \cdot \text{number of visits}_{ij} + \beta_{6,\text{LRTI}} \cdot \text{year of diagnosis}_{ij} + \tau_{i,\text{LRTI}}$$

Gastric Acid Related (GAR):

$$\eta_{ij,\text{GAR}} = \beta_{0,\text{GAR}} + \beta_{1,\text{GAR}} \cdot \text{age group}_{ij} + \beta_{2,\text{GAR}} \cdot \text{education level}_{ij} + \beta_{3,\text{GAR}} \cdot \text{wealth status}_{ij} + \beta_{4,\text{GAR}} \cdot \text{household size}_{ij} + \beta_{5,\text{GAR}} \cdot \text{number of visits}_{ij} + \beta_{6,\text{GAR}} \cdot \text{year of diagnosis}_{ij} + \tau_{i,\text{GAR}}$$

Urinary Track Infection (UTI):

$$\eta_{ij,\text{UTI}} = \beta_{0,\text{UTI}} + \beta_{1,\text{UTI}} \cdot \text{age group}_{ij} + \beta_{2,\text{UTI}} \cdot \text{education level}_{ij} + \beta_{3,\text{UTI}} \cdot \text{wealth status}_{ij} + \beta_{4,\text{UTI}} \cdot \text{household size}_{ij} + \beta_{5,\text{UTI}} \cdot \text{number of visits}_{ij} + \beta_{6,\text{UTI}} \cdot \text{year of diagnosis}_{ij} + \tau_{i,\text{UTI}}$$

Each coefficient ($\beta_{m,k}$) (for $m = 0, 1, \dots, 6$) represents the effect of the corresponding predictor variable on the log-odds of diagnosis k relative to Coryza. The conditional probability of each diagnosis category (other than Coryza) is given by:

Malaria:

$$P(Y_{ij} = \text{Malaria}) = \frac{\exp(\eta_{ij,\text{Malaria}})}{1 + \exp(\eta_{ij,\text{Malaria}}) + \exp(\eta_{ij,\text{LRTI}}) + \exp(\eta_{ij,\text{GAR}}) + \exp(\eta_{ij,\text{UTI}})}$$

Lower Respiratory Tract Infection (LRTI):

$$P(Y_{ij} = \text{LRTI}) = \frac{\exp(\eta_{ij,\text{LRTI}})}{1 + \exp(\eta_{ij,\text{Malaria}}) + \exp(\eta_{ij,\text{LRTI}}) + \exp(\eta_{ij,\text{GAR}}) + \exp(\eta_{ij,\text{UTI}})}$$

Gastric Acid Related (GAR):

$$P(Y_{ij} = \text{GAR}) = \frac{\exp(\eta_{ij,\text{GAR}})}{1 + \exp(\eta_{ij,\text{Malaria}}) + \exp(\eta_{ij,\text{LRTI}}) + \exp(\eta_{ij,\text{GAR}}) + \exp(\eta_{ij,\text{UTI}})}$$

Urinary Track Infection (UTI):

$$P(Y_{ij} = \text{UTI}) = \frac{\exp(\eta_{ij,\text{UTI}})}{1 + \exp(\eta_{ij,\text{Malaria}}) + \exp(\eta_{ij,\text{LRTI}}) + \exp(\eta_{ij,\text{GAR}}) + \exp(\eta_{ij,\text{UTI}})}$$

For the reference category (**Coryza**), the probability is:

$$P(Y_{ij} = \text{Coryza}) = \frac{1}{1 + \exp(\eta_{ij,\text{Malaria}}) + \exp(\eta_{ij,\text{LRTI}}) + \exp(\eta_{ij,\text{GAR}}) + \exp(\eta_{ij,\text{UTI}})}$$

All analyses were performed using R software. The R package *brms*² was used to fit the model.

Supplemental Material 2: Conditional effect plot

Fig. 1 shows the expected probability of being diagnosed with a particular disease conditioned on a particular level of a predictor of interest - the conditional effect plots. Furthermore, it shows how the predicted probabilities differ across various levels of the predictor. Other than age groups and number of visits, in which the probability of being diagnosed with Malaria is higher in children aged 0-14 years and those patients who had one visit, respectively, the expected probability of being diagnosed with a particular condition remained relatively the same across different levels of the predictor of interest. Furthermore, the probability of being diagnosed with coryza is higher in older patients. On the other hand, the probability of being diagnosed with lower respiratory tract infection has decreased since 2018. This is not the case with Malaria.

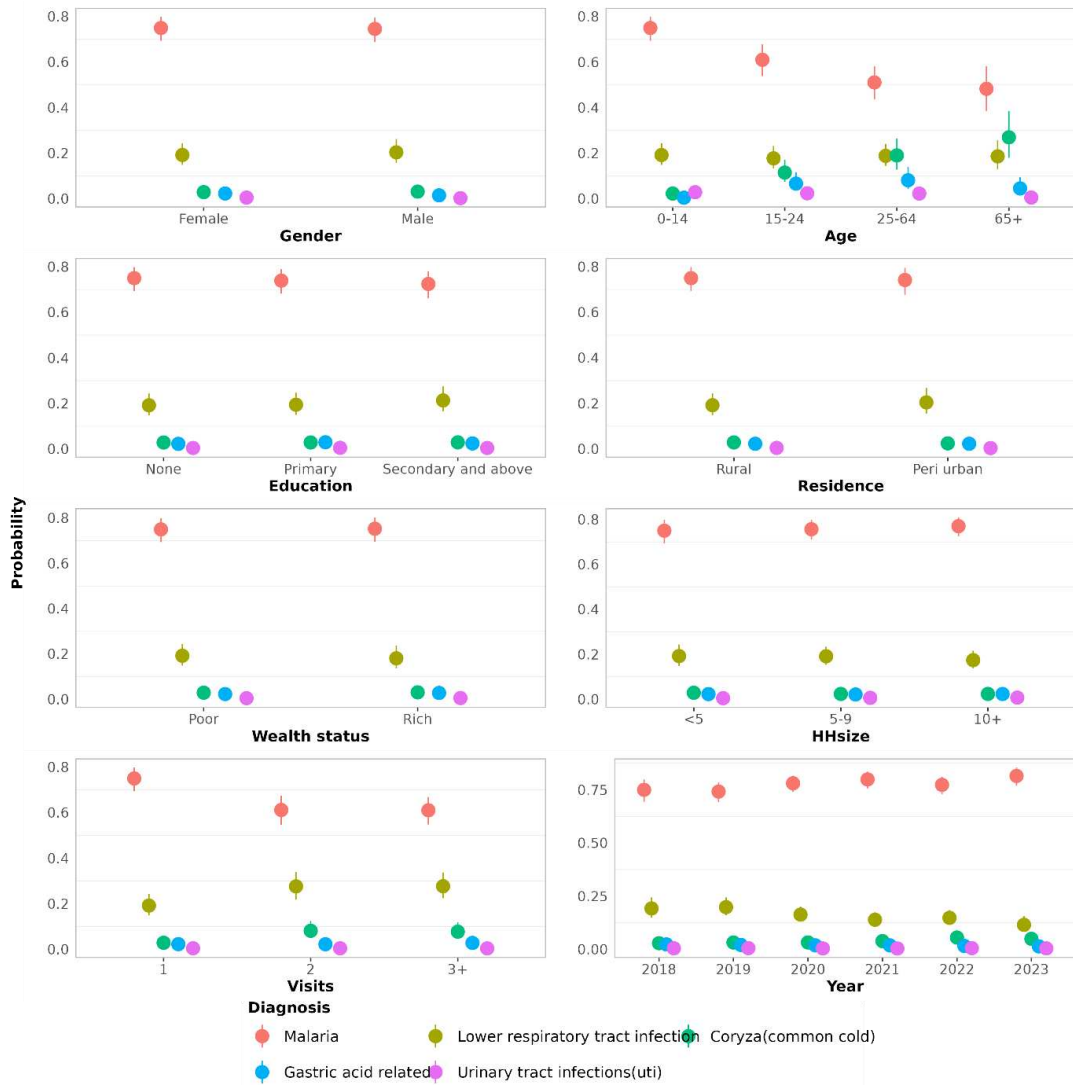


Fig. 1: A comparison of the conditional effect plots for all predictors.

References

1. Kazembe LN, Namangale JJ. A Bayesian multinomial model to analyse spatial patterns of childhood co-morbidity in Malawi. *Eur J Epidemiol* 2007;22(8):545–56.
2. Bürkner P-C. **brms** : An R Package for Bayesian Multilevel Models Using *Stan*. *J Stat Softw* [Internet] 2017 [cited 2023 Nov 23];80(1). Available from: <http://www.jstatsoft.org/v80/i01/>