**S2 Bayesian model formulation.**

Let **Y***jt* be the average number of deaths (all-cause or malaria-specific) in village *j* at time interval *t*. We assume that **Y***jt* arises from a negative binomial distribution.

 $Y\_{jt}\~dnegbin(P\_{jt},r)$

where P*jt*, is the proportion of deaths occurring in village *j* at time interval *t* and *r*is the dispersion parameter with,

 $ μ\_{j}=r\frac{1-p}{p}$ and$ σ\_{i}^{2}=r(1-p)p^{-2}$.

We modeled the association above between covariates (*X*) and mortality status of individuals by village on the logit, as

$logit\left(μ\_{jt}\right)=logit\left(N\_{jt}\right)+β\_{0}+\sum\_{1}^{k}β\_{u}X\_{u}+ϕ\_{j}+ε\_{t}$ , *u=1,2,……k*

where $μ\_{jt}$ is the number of deaths in each village at time *t*, $N\_{jt} $the total person time contributed by persons in each village as discrete months, $β\_{i}$ the regression coefficients, $ϕ\_{j}$ the village specific spatial effects and $ε\_{t}$ the temporal (monthly) random effects.

We assumed that $ϕ\_{j}$ are parameters from a latent spatial process modelled by a Gaussian distribution with covariance matrix quantifying the relation between any pair of villages as a function of their distance irrespective of the direction using an exponential correlation function, that is $∅ \~ MVN(0,∑)$, $∑\_{kl}= σ\_{1}^{2} exp⁡(-ρd\_{kl})$ where $σ\_{1}^{2}$ is the spatial variation, $d\_{kl}$ is the distance between villages *k* and *l*, and $ρ$ is the rate of correlation decay with increasing distance. The minimum distance at which the spatial variation is less than 5% is called range and can be obtained from the value $3/ρ$ (49). Temporal effect ($ε\_{t}$) was modeled by an autoregressive process of order 2. We specified non-informative normal prior distributions with mean zero and large variance for the *βi* *i*=1,…… regression coefficients, an inverse gamma prior for $r \~ IG(1.01, 0.001)$, an inverse gamma priors for $σ\_{e}^{2}$ and $σ^{2}$. A gamma prior for $ρ$, that is $σ\_{e}^{2} , σ^{2} \~ IG(2.01, 1.01)$ and $ρ \~ G(0.01, 0.01)$.

The model was fitted using Markov Chain Monte Carlo (MCMC) simulation algorithm in OpenBugs version 3.1.2 (Imperial College and Medical Council, London, UK) to estimate model parameters (50). We ran a single chain sampler discarding the first 10,000 iterations. Convergence was assessed by Gelman-Rubin diagnostic (51) and attained at 100,000 iterations.