

RESEARCH

Spatial Models for Infants HIV/AIDS Incidence Using an Integrated Nested Laplace Approximation Approach

Susan N Mutua^{1,3*†} and Thomas N O Achia^{2,3}

*Correspondence:

nzulasue@gmail.com

¹Kenya National AIDS & STI Control Programme, Afya Annex, KNH Grounds, Nairobi, Kenya

Full list of author information is available at the end of the article

[†]Equal contributor

Abstract

Background: Kenya has made significant progress in the elimination of mother to child transmission of HIV through increasing access to HIV treatment and improving the health and well-being of women and children living with HIV. Despite this progress, broad geographical inequalities in infant HIV outcomes still exist. This study aimed at assessing the spatial distribution of HIV amongst infants, areas of abnormally high risk and associated risk factors for mother to child transmission of HIV using INLA and SPDE approach.

Methods: Data were obtained from the Early infant diagnosis (EID) database that is routinely collected for infants under one year for the year 2017. We performed both areal and point-reference analysis. Bayesian hierarchical Poisson models with spatially structured random effects were fitted to the data to examine the effects of the covariates on infant HIV risk. Spatial random effects were modelled using Conditional autoregressive model (CAR) and stochastic partial differential equations (SPDEs). Inference was done using Integrated Nested Laplace Approximation. Posterior probabilities for exceedance were produced to assess areas where the risk exceeds 1. The Deviance Information Criteria (DIC) selection was used for model comparison and selection.

Results: CAR model outperformed similar competing models in modeling and mapping HIV Relative Risk in Kenya. It had a smaller DIC among the rest (DIC = 306.36) The SPDE model outperformed the spatial GLM model based on the DIC statistic. Highly active antiretroviral therapy (HAART) and breastfeeding were found to be negatively and positively associated with infant HIV positivity respectively [-0.125, 95% Credible Interval (Cred. Int.) = -0.348, -0.102], [0.178, 95% Cred. Int. -0.051, 0.412].

Conclusion: The study provides relevant strategic information required to make investment decisions for targeted high impact interventions to reduce HIV infections among infants in Kenya.

Keywords: Markov Chain Monte Carlo; Integrated Nested Laplace Approximation; Stochastic Partial Differential Equation; Early Infant Diagnosis; HIV Exposed Infant; Gaussian Markov Random Field

Background

Globally, there has been a remarkable progress in elimination of mother to child HIV transmission. There was a decline in the number of new infections among children from 290,000 (250,000-350,000) in 2010 to 150,000 (110,000-190,000) in 2015, reflecting the scale-up of coverage of prevention of mother-to-child transmission services (PMTCT) [1]. In Sub-Saharan countries, new HIV paediatric infections declined marginally between 2010 to 2015. A decline of 66% was reported in Southern and Eastern Africa while 31% was reported in Western and central Africa [1].

Despite these achievements, concerted efforts are required to reach the UNAIDS targets by 2020 [2]. This can be achieved through ensuring all pregnant women living with HIV receive lifelong ART, scale up the provision of early diagnosis, treatment optimization and care for HIV exposed infants to prevent mother to child transmission of HIV [3]. Elimination of mother to child transmission (eMTCT) of HIV would directly contribute to the attainment of Sustainable Development Goal targets of reduction of maternal mortality ratio, ending preventable deaths of newborns and children under 5 years and ending the AIDS epidemic [4].

Kenya is one of the 22 priority countries focused for reduction of mother-to-child Transmission (MTCT) of HIV and has been chosen to be validated for the pre-elimination of MTCT of HIV and Syphilis by 2021. An estimated 6,613 infants were infected with HIV through mother to child transmission in 2015 [5]. This is a decline in mother to child transmission rates from 16% in 2013 to 8.3% in 2015 [5]. These gains are attributed to the increase in HIV treatment coverage for women and infants.

The government of Kenya through the Ministry of Health (MOH) established the Prevention of mother to child transmission of HIV (PMTCT) program as part of continuing strategies for dealing with the epidemic and interventions to reduce MTCT of HIV. The goal of the program is to reduce the rate of MTCT of HIV to less than 5% and reduce maternal mortality by 50%.

In Kenya, the early infant diagnosis (EID) program, under the umbrella of the PMTCT program, is responsible for making HIV diagnosis in HIV exposed infants (HEI) and young children under 18 months of age. Polymerase chain reaction (PCR) is the most common virological test used in PMTCT settings for HIV exposed babies and there are seven laboratories nationally with capacity to conduct PCR testing of HIV.

Despite the marked progress in elimination of mother to child transmission (eMTCT) of HIV, few counties in Kenya are still registering high number of infections. New HIV infections amongst infants exhibits marked geographical disparities with counties contributing disproportionately high number of new infections annually. To accelerate and achieve a drastic reduction in the new HIV infections, focussed efforts needs to be directed to the non performing counties.

Disease mapping has developed immensely in the recent years due to availability of geo-referenced data with advances in computing, geographical information systems (GIS) and statistical methodology. The interest lies in providing estimates of the relative risks of a disease across a geographical study area, assessing clustering and clusters of disease and assessing geographical distribution of disease in relation to

potential risk factors (spatial regression). Bayesian methods which offer a flexible and robust approach are increasingly being utilized in disease mapping.

In the last few years, MCMC methods have boosted the implementation of fixed effects and hierarchical models particularly in spatial and spatial temporal field [6]. Despite the progress made in Bayesian computing, Markov Chain Monte Carlo (MCMC) methods are not without potential problems. It's samplers involve computationally and time intensive simulations especially for high dimensional models such as hierarchical models. Furthermore, Parameter estimation might be impossible and the algorithms may induce large Monte Carlo standard errors if they're not run for many iterations [7]. Additionally, MCMC methods present issues with convergence of the algorithm to the posterior distribution as well as choice of prior distributions.

Analysis of large datasets with vast level of spatial disaggregation could lead to long computation time to perform Bayesian inference via MCMC. Cross validation tests and sensitivity analyses might be impractical because of the computational demands that come with MCMC. This can translate to poor interpretation of the results.

The Integrated Nested Laplace Approximation (INLA) proposed by Havard Rue [8] provides a alternative approach to handling these complexities and providing precise and consistent estimates within a short computational time.

INLA is designed for latent Gaussian models ranging from generalised linear (mixed) models, generalised additive (mixed) models, geoadditive models and time series models. The development of R package named R-INLA has proved valuable for the implementation of INLA. INLA can be integrated with the Stochastic Partial Differential Equation (SPDE) approach proposed by [9] to execute spatial and spatio-temporal models for geostatistical data.

There has been a gradual increase in the utilization of these methods in analysis of epidemiological and public health data particularly in spatial and spatial temporal models.

This paper utilizes spatial models to assess the association between maternal and infants covariates and HIV sero-conversion among HEI using INLA.

Methods

Study area and Data

The study was conducted in Kenya. The country is divided into 47 administrative units (counties). We used 2017 early infant diagnosis (EID) program data collected routinely by the Ministry of Health to investigate the spatial patterns of HIV amongst infants in Kenya. The covariates considered were maternal prophylaxis and infant breastfeeding. HEI are tested through EID which gives an opportunity for early identification of HIV and linkage to care and treatment services. The facility-level data comprised of 68,600 PCR tests types (Initial, 2nd, 3rd and confirmatory), PCR test results, sex and age of infant, infant and maternal prophylaxis, mother HIV status, infant breastfed, Entry point, testing laboratory, date samples received, tested and dispatched collected in the 47 counties in Kenya. The analyses were restricted to infants under one year born to HIV positive women. The data were aggregated to provide county level summaries that were used for areal analysis. The facility level EID data was used for geostatistical analysis.

Statistical analysis

Variables

The study aimed to find out the relationship between infant HIV count test results and important covariates consisting of Highly active antiretroviral therapy (HAART) and proportion of infants breastfed. We excluded infant prophylaxis since it is highly correlated with HAART.

Statistical models

We developed Bayesian hierarchical Poisson regression models to investigate the spatial heterogeneity of HIV across the counties and assess the effects of the selected covariates.

For the i^{th} district, the count of HIV cases y_i is modelled as $y_i \sim \text{Poisson}(E_i\theta_i)$, where the mean θ_i is the risk of infection and E_i is the expected number of infants infected with HIV in district i , $i = 1, 2, \dots, 47$.

The linear predictor is defined as:

$$\eta_i = \beta_0 + \sum_{m=1}^M \beta_m t_{mi} + \sum_{l=1}^L f_l z_{li} + \varepsilon_i, \quad (1)$$

where β_0 denotes HIV outcome rate for all the 47 counties, β_m quantify the effect of the covariates $t = (t_1, \dots, t_m)$ and $f_l(\cdot)$ are the spatial random effects that capture spatial variation modelled using conditional intrinsic autoregressive model (ICAR) prior distribution[10]. The ICAR model is given as

$$\phi_i | \phi_{-i} \sim \text{Normal} \left(\frac{1}{n_i} \sum_{j \sim i} \phi_j, \frac{\sigma^2}{n_i} \right). \quad (2)$$

The conditional expectation of random effect ϕ_i is the average of the effects of its neighbours. The conditional variance depends on its number of neighbours n_i . An area with many neighbours will have a smaller variance.

To investigate the association between the outcome variable and the covariates, we fitted four Poisson models to the county level data :

- 1 **Model 1:** The ordinary Poisson regression model.
- 2 **Model 2:** The generalized linear mixed model spatially unstructured random effects.
- 3 **Model 3:** The generalized linear mixed model spatially structured random effects.
- 4 **Model 4:** The generalized linear mixed model with both spatially unstructured and structured random effects.

For the geostatistical data, we investigated spatial dependency in the Pearson residuals by making a sample variogram. We proceeded with fitting two models:

- 1 **Model 1:** The spatial component model (GLM Poisson model).
- 2 **Model 2:** The model without spatial component (SPDE model).

Bayesian inference was done in R software [11] using Integrated Nested Laplace approximation (R-INLA) package. A stochastic partial differential equation (SPDE) with INLA was employed to estimate the posterior marginals in the geostatistical analysis.

Model comparison and selection was done using the deviance information criterion (DIC) that takes into account the trade off between model fit and complexity [12]. The best model was given by the model with the smallest value of DIC.

In the SPDE model, the Gaussian random field with its covariance function was represented as a Gaussian Markov random field. A non-stationary model was achieved by modifying the SPDE to obtain the Gaussian Markov random field with a defined dependence structure that is different from the stationary Mtern covariance. The local nature of differential operators was used to allow for local specification of the range and variance parameters.

Results

Summary Results

Table 1 presents the summary statistics of the variables in the data. In total, of the 68,600 HIV exposed infants who provided their blood samples for PCR HIV testing, 2,363 (3.4%) turned positive for HIV. Of the 2,363 HIV positive women, 86.33% breastfed their infants while the remaining 13.24% did not breastfeed their infants. Majority (77.06%) of these women accessed the health facility through the MCH/PMTCT. 61.32% of the infants who turned out positive were on prophylaxis while 63.98% of their mothers were on maternal prophylaxis. 81.17% of the infants who turned HIV positive had their first diagnosis after 2 months.

Areal data analysis

The results of fitting the hierarchical models are shown in table 2. Based on these results, we deduce that the model with spatially structured random effects (Model 3) offered a better fit (DIC 306.36).

Increased use of HAART is associated with mother to child transmission of HIV. -0.802 (95% credible interval:-2.19, 0.60). Breast feeding is positively associated with mother to child transmission of HIV 0.558 (95% credible interval: -1.622,-2.778). The results however suggest that the covariates are not significant. Figure 1 shows the relative risk map of the counties based on the best fitting model. The relative risk ranges from 0.71 to 1.71. Counties that were predicted to have high risk of Infant HIV infection were Makueni, Lamu, Turkana, Marsabit, Samburu, Mombasa, Nairobi, Narok, Taita Taveta, Kilifi. Bungoma, Embu, Muranga, Garissa, Tana river, Elgeiyo Marakwet, Homabay and Nyandarua were associated with lower risk of infant HIV infection ($RR < 0.8$). The exceedence probability map in Figure 2 displays counties with relative risk above the national risk ($RR > 1$). This map confirms counties associated with high risk of HIV infection (darker color). The distribution of spatial random effects (Fig 3) revealed strong spatial patterns at multiple scales. According to these patterns, the risk of HIV infection is associated with living in the following counties; Makueni, Turkana, Marsabit, Lamu. This map

shows a similar pattern to the relative risk map signifying counties highly affected by spatially structured random effects.

Clustering of risk and elevated risk can be observed in the North-west and south west counties of Kenya.

Geostatistical analysis

The variogram (Figure 4) obtained from the Poisson GLM indicates presence of spatial correlation in the data. Table 3 displays the posterior estimates of the Poisson GLM and the SPDE model. The best fitting model was the SPDE model (DIC 5591.732). Consistent with the results of the areal analysis, HAART was found to be negatively associated with infant positivity whereas breastfeeding was positively associated with infant positivity. The associations were however not significant.

Figure 5 shows a map of the sampled locations and the selected mesh is shown in Figure 6. The triangulation produced a mesh with 1249 vertices. To avoid the boundary effects, we used a mesh that extends the study region. The choice of mesh is a trade-off between the GMRF representation accuracy and the computational costs [13].

The map of the spatial field (Figure 7 top left) reveals that the spatial random effects causes an increase or decrease in the expected disease counts in specific regions.

The spatial pattern of the posterior mean of the latent field and of the mean response are similar whereas the latent field has a higher variability than the mean response.

Discussion

In this work, deterministic Bayesian approaches (INLA and INLA-SPDE) were used for analysis of areal and point reference data to assess the spatial distribution of HIV amongst infants, risk factors for mother to child transmission and to produce continuous maps of the positivity rates. We employed INLA to fit Bayesian hierarchical spatial models within the R library INLA. CAR model was identified to be suitable for modeling and mapping relative risk of HIV amongst infants in Kenya. The findings revealed that breastfeeding increases mother to child transmission of HIV. Previous studies [14, 15], noted a substantially higher risk of infection among breast fed infants within the first months of breastfeeding compared to later months. A high early transmission rate might be explained by the milk which is rich in HIV infected cells and the immaturity of the infant's immune system [16]. In resource limited settings complementary feeding increases the risk of morbidity and mortality from infectious disease. Until recently, WHO recommended that HIV positive mothers breastfeed exclusively for the first six months and continue breastfeeding with appropriate complementary foods for at least 12 months while taking their antiretroviral to reduce risk of post-natal transmission. [17] showed that perinatal transmission can be greatly reduced in breastfeeding populations on antiretroviral therapy.

The results of the study indicated that the use of HAART by the mother reduces the

transmission rate of HIV. The use of HAART clearly stands out as a key determinant of MTCT risk as has been consistently reported in many studies [18–20]. The risk of HIV infection among infants was found to be high in Makueni, Lamu, Turkana, Marsabit, Samburu, Mombasa, Nairobi, Narok, Taita Taveta, Kilifi counties suggesting the need of geographical prioritization of infant HIV prevention interventions to optimise reduction of new HIV infections. The random effects map revealed residual variation suggesting unaccounted variation after including the covariates in the model. Additionally, the spatial patterns in the random effects suggested possible covariates omitted from the model. INLA which is a promising alternative to the MCMC is a recent methodology for Bayesian inference used in hierarchical models. The numerical inference approach has gained popularity due to the fast and accurate estimates to the posterior distributions produced by the methodology. This alleviates one of the most important bottlenecks associated with MCMC which is computationally intensive especially for models that are complex in nature.

We employed INLA-SPDE approach to perform Bayesian inference on spatial hierarchical Gaussian fields. This model was applied to produce continuous predictions of positivity rates. The SPDE model was found to be suitable for modelling and predicting point reference data as compared to the Poisson GLM. The posterior covariates effects were similar to those of the areal analysis. It is apparent from the maps of the random effects and the spatial random field that inclusion of spatial random effects yields strong latent spatial patterns that were not explained by the explanatory variables. As shown in Figure 7, the predicted risk of infant HIV infection is high in the western part of Kenya and lowest in the eastern regions. The possible reason for this could be that the data exhibits clustering in the western regions with little or no observations in the other parts. The standard errors are high on the peripherals which should be expected since they are extensions of the triangulation. One advantage of INLA-SPDE is its capability of estimation and prediction of the posterior marginal distributions of the model parameters and the model responses without carrying out extensive simulations. A major limitation with INLA-SPDE is that it becomes computationally intensive when dealing with non-Gaussian likelihood [9].

Limitations of this study include the use of routinely collected data which may not be of quality due to missing information. Additionally, the data has limited information on potential explanatory variables that could be important in explaining the positivity rates. Lack of data on important covariates such as virological status and maternal immunologic which could influence transmission rates.

Conclusion

The aim of this study was to investigate the geographical variation of HIV and identify risk factors among infants in Kenya in 2017. INLA which is a computationally effective alternative to MCMC was used for Bayesian inference. This study has shown the determinants of infant HIV infection and the spatial distribution through mapping of HIV risk. It is apparent that there are still geographical disparities in infant HIV infection and targeted PMTCT interventions and resources should be directed to counties that have high infection rates. It is hoped that the presented outcome variations will stimulate further research efforts to investigate the reasons

underlying the disparities and inform policy. Many aspects of this study could be extended. Further research is required to determine suitable spatio-temporal models for modeling and mapping relative risk of HIV among infants in Kenya. This will allow to explain evolution of the relative risk of HIV in space and time. Incorporating other covariates for mother to child transmission of HIV that were not captured in the analysis could yield important insights on the spatial distribution of the disease.

Abbreviations

PMTCT- Prevention of Mother to Child Transmission
 EID - Early Infant Diagnosis
 MTCT - Mother to Child Transmission
 PCR - Polymerase Chain Reaction
 INLA - Integrated Nested Laplace Approximation
 MCMC - Markov Chain Monte Carlo
 SPDE - Stochastic Partial Differential Equations
 HAART - Highly Active Antiretroviral Therapy

Competing interests

The authors declare that they have no competing interests.

Author's contributions

SNM conceived and designed the study, analyzed the data, prepared figures and tables, drafted the manuscript. TNOA analyzed the data, authored or reviewed drafts of the paper, trained the lead author in analytical methods, supervised the writing of the manuscript.

Ethics approval and consent to participate

Ethical approval for the study was obtained from Strathmore University Institutional Ethics Review Committee (SU-IERC).

Acknowledgements

We acknowledge NASCOP, Kenya Ministry of Health for letting us use their data with minimal restrictions.

Author details

¹Kenya National AIDS & STI Control Programme, Afya Annex, KNH Grounds, Nairobi, Kenya. ²School Mathematics, Statistics and Computer Science, University of KwaZulu Natal, Scottsville, 3209 Pietermaritzburg, South Africa. ³Strathmore University, Ole Sangale Road, Nairobi, Kenya.

References

1. Joint United Nations Programme on HIV/AIDS: Prevention gap report, 266–267 (2016)
2. PEPFAR, UNAIDS and UNICEF and others: Who. start free. stay free. aids free. A Super-Fast-Track Framework for Ending AIDS Among Children, Adolescents and Young Women by 2020 (2016)
3. Mother-to-child Transmission of HIV. http://who.int/hiv/topics/mtct/about/en_links.html
4. Griggs, David and Stafford-Smith, Mark and Gaffney, Owen and Rockström, Johan and Öhman, Marcus C and Shyamsundar, Priya and Steffen, Will and Glaser, Gisbert and Kanie, Norichika and Noble, Ian: Policy: Sustainable development goals for people and planet. *Nature* **495**(7441), 305 (2013)
5. Ministry of Health, National AIDS and STI Control Programme: Kenya Framework for Elimination of Mother to Child Transmission of HIV 2016-2021, (2015)
6. Ntzoufras, Ioannis: Bayesian Modeling Using WinBUGS vol. 698. John Wiley & Sons, Hoboken, New Jersey (2011)
7. Gilks, Walter R and Richardson, Sylvia and Spiegelhalter, David: Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC, New York (1995)
8. Rue, Håvard and Martino, Sara and Chopin, Nicolas: Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)* **71**(2), 319–392 (2009)
9. Lindgren, Finn and Rue, Håvard and Lindström, Johan: An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**(4), 423–498 (2011)
10. Besag, Julian and Kooperberg, Charles: On conditional and intrinsic autoregressions. *Biometrika* **82**(4), 733–746 (1995)
11. Team, R Core and others: R: A language and environment for statistical computing (2013)
12. Spiegelhalter, David J and Best, Nicola G and Carlin, Bradley P and Van Der Linde, Angelika: Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4), 583–639 (2002)
13. Blangiardo, Marta and Cameletti, Michela: Spatial and Spatio-temporal Bayesian Models with R-INLA. John Wiley & Sons, Hoboken, New Jersey (2015)
14. Nduati, Ruth and John, Grace and Mbori-Ngacha, Dorothy and Richardson, Barbra and Overbaugh, Julie and Mwatha, Anthony and Ndinya-Achola, Jeckoniah and Bwayo, Job and Onyango, Francis E and Hughes, James

- and others: Effect of breastfeeding and formula feeding on transmission of hiv-1: a randomized clinical trial. *Jama* **283**(9), 1167–1174 (2000)
15. Colebunders, Robert and Kapita, Bila and Nekwei, Wasadidi and Bahwe, Yemvula and Lebughe, Izaley and Oxtoby, Margaret and Ryder, Robert: Breastfeeding and transmission of hiv. *Lancet* **2**(8626-8627), 1487 (1988)
 16. Southern, SO: Milk-borne transmission of hiv. characterization of productively infected cells in breast milk and interactions between milk and saliva. *Journal of human virology* **1**(5), 328–337 (1998)
 17. Slater, Mackenzie and Stringer, Elizabeth M and Stringer, Jeffrey SA: Breastfeeding in hiv-positive women. *Pediatric Drugs* **12**(1), 1–9 (2010)
 18. Townsend, Claire L and Cortina-Borja, Mario and Peckham, Catherine S and de Ruiter, Annemiek and Lyall, Hermione and Tookey, Pat A: Low rates of mother-to-child transmission of hiv following effective pregnancy interventions in the united kingdom and ireland, 2000–2006. *Aids* **22**(8), 973–981 (2008)
 19. Ledergerber, Bruno and Egger, Matthias and Opravil, Milos and Telenti, Amalio and Hirschel, Bernard and Bategay, Manuel and Vernazza, Pietro and Sudre, Philippe and Flepp, Markus and Furrer, Hansjakob and others: Clinical progression and virological failure on highly active antiretroviral therapy in hiv-1 patients: a prospective cohort study. *The Lancet* **353**(9156), 863–868 (1999)
 20. Becquet, Renaud and Ekouevi, Didier K and Arrive, Elise and Stringer, Jeffrey SA and Meda, Nicolas and Chaix, Marie-Laure and Treluyer, Jean-Marc and Leroy, Valérie and Rouzioux, Christine and Blanche, Stéphane and others: Universal antiretroviral therapy for pregnant and breast-feeding hiv-1-infected women: towards the elimination of mother-to-child transmission of hiv-1 in resource-limited settings. *Clinical infectious diseases* **49**(12), 1936–1945 (2009)

Figures

Figure 1 Infants HIV Relative risk in Kenya in 2017 Figure showing infants relative risk in Kenya. Counties with darker colors depict higher Relative risk

Figure 2 Posterior relative risk exceedance probability map $p(\xi > 1|y)$ This map depicts the areas where it is more likely to observe Relative Risk higher than the threshold

Figure 3 Estimates of the random effects from the spatially structured model

Figure 4 Sample variogram of the pearsons residuals of the Poisson GLM

Figure 5 Locations of sampled facilities

Figure 6 Mesh with sampled locations The mesh used extended beyond the shape file and it produced 1249 vertices

Figure 7 Spatial random field (top) and the predicted mean response (bottom)

Tables

Table 1 Summary of infant and maternal risk factors associated with mother to child transmission of HIV in Kenya 2017

Variable	Total (N = 68,600)	HIV Positive (N = 2,363)
Entry Point		
CCC/PSC	6,544 (9.53)	233 (9.86)
IPD	293 (0.42)	87 (3.68)
Maternity	409 (0.59)	18 (0.76)
MCH/PMTCT	59,816 (87.19)	1,821 (77.06)
OPD	774 (1.12)	145 (6.13)
Other	350 (0.51)	35 (1.48)
Infant breastfed		
BF	57,889 (84.39)	2,040 (86.33)
NBF	10,711 (15.61)	313 (13.24)
Sex		
Female	34,651 (50.51)	1,197 (50.65)
Male	33,188 (48.38)	1,132 (47.91)
Infant prophylaxis		
Yes	51,681 (75.33)	1,449 (61.32)
No	16,919 (24.66)	914 (38.67)
Maternal prophylaxis		
Yes	62,769 (91.57)	1,512 (63.98)
No	5,831 (8.50)	851 (36.02)
Age at diagnosis		
Under 8 weeks	14255 (20.78)	445 (18.83)
Over 8 weeks	54345 (79.22)	1918 (81.17)

Table 2 Posterior statistics of the four models

Variable	Model 1	Model 2	Model 3	Model 4
β_0	3.3285 (1.8445, 4.7958)	3.9853 (1.2337, 6.6126)	3.0724 (0.3483, - 5.6721)	3.7619 (1.0254, 6.3732)
HAART	-0.3422 (-1.1703, 0.4926)	-0.5565 (-2.0945, 0.9918)	-0.8020 (-2.1955, 0.6030)	-0.6588 (-2.1714, 0.8674)
BF	-0.4925 (-1.7174, -0.7381)	-0.1736 (-2.4892, -2.1779)	0.5576 (-1.6219, -2.7776)	0.0399 (-2.2514, -2.3716)
medianage	-0.3028 (-0.4115, -0.1922)	-0.3777 (-0.5661, -0.1791)	-0.3173 (-0.5125 , -0.1114)	-0.3632 (-0.5504, -0.1660)
τ_v		20.47 (9.400, 39.498)		22.754 (10.084, 45.226)
τ_u			4.703 (2.134, 9.079)	8622.668 (142.832, 39713.666)
DIC	393.11	307.39	306.36	307.67
pD	4.031	27.14	27.55	26.93

Table 3 Posterior statistics of the Poisson GLM and Poisson GLM + SPDE

Variable	Poisson GLM	Poisson GLM + SPDE
β_0	-0.059 (-0.27727402, 0.15261972)	-0.609 (-0.904, -0.322)
HAART	-0.129 (-0.313, 0.058)	-0.125 (-0.348, 0.102)
BF	0.135178 (-0.062, 0.338)	0.178 (-0.051, 0.412)
DIC	8463.834	5591.732
σ		0.019 (0.000, 0.003)
κ		0.495 (0.008, 1.388)
Range		14.621 (0.623, 48.579)