Full Title: Bayesian hierarchical modeling of joint spatiotemporal risk patterns for Human Immunodeficiency Virus (HIV) and Tuberculosis (TB) in Kenya

Short Title: Spatiotemporal modeling of joint risk patterns of HIV and TB in Kenya

Verrah A. Otiende[1*], Thomas N. Achia[2], Henry G. Mwambi[2]

[1]Department of Mathematical Sciences, Pan African University Institute of Basic Sciences Technology and Innovation, Nairobi, Kenya

[2]School of Mathematics, Statistics & Computer Science, University of KwaZulu Natal, Pietermaritzburg, South Africa

[*]Corresponding author

Email: otiende.verrah@students.jkuat.ac.ke; verrahodhiambo@gmail.com

VO, TO and HM conceived and designed the study. VO wrote the first draft of the manuscript. VO and TO analyzed data. HM contributed to data analysis. All authors contributed to reviewing literature, interpretation of the results and writing of the manuscript. All authors read and approved the final manuscript.

Abstract

The spatiotemporal modeling of multiple diseases simultaneously is a recent extension that advances the space-time analysis to model multiple related diseases simultaneously. This approach strengthens inferences by borrowing information between related diseases. Numerous research contributions to spatiotemporal modeling approaches exhibit their strengths differently with increasing complexity. However, contributions that combine spatiotemporal approaches to modeling of multiple diseases simultaneously are not so common. We present a full Bayesian hierarchical spatio-temporal approach to the joint modeling of Human Immunodeficiency Virus and Tuberculosis incidences in Kenya. Using case notification data for the period 2012 - 2017, we estimated the model parameters and determined the joint spatial patterns and temporal variations. Our model included specific and shared spatial and temporal effects. The specific random effects allowed for departures from the shared patterns for the different diseases. The space-time interaction term characterized the underlying spatial patterns with every temporal fluctuation. We assumed the shared random effects to be the structured effects and the disease-specific random effects to be unstructured effects. We detected the spatial congruence in the distribution of Tuberculosis and Human Immunodeficiency Virus in approximately 29 counties around the western, central and southern regions of Kenya. The distribution of the shared relative risks had minimal difference with the Human Immunodeficiency Virus disease-specific relative risk whereas that of Tuberculosis presented many more counties as high-risk areas. The flexibility and informative outputs of Bayesian Hierarchical Models enabled us to identify the similarities and differences in the distribution of the relative risks associated with each disease. Estimating the Human Immunodeficiency Virus and Tuberculosis shared relative risks provide additional insights towards collaborative monitoring of the diseases and control efforts.

Introduction

Disease mapping as a modeling approach is commonly used to describe the geographical distribution of disease burden thereby generating hypotheses on their possible causes and differences [1]. Statistical methods for characterization of the geographical variations have contributed significantly towards advancing the focus of disease mapping to incorporate other analysis techniques. Including the time dimension has given rise to the spatiotemporal modeling of the variation of disease risk which enables simultaneously studying the spatial patterns and temporal variations thereby giving deeper insights over purely spatial mapping [2,3]. Another extension is the joint spatial modeling of multiple related diseases with common risk factors, which outspread the standard univariate disease mapping methodologies. It generates both the shared and divergent trends thereby increasing the estimation precision of disease risk [4]. The latest and most recent extension is the joint spatiotemporal modeling of multiple diseases. This approach advances the space-time analysis to model multiple diseases simultaneously thereby strengthening inference by borrowing information across neighboring regions and between related diseases or sub-populations with common aetiological factors [5]. The benefits of borrowing information lie in the ability to observe concurrency of patterns and to allow conditioning of one disease on others [6] which is very valuable when accounting for uncertainty due to sparse disease count or underreporting [7]

Studies combining spatiotemporal approaches to modeling multiple diseases simultaneously are not so common despite the development and application of novel computational techniques. Reviews from literature show few contributions of these approaches that exhibit their strengths differently with increasing complexities. [8] used a Bayesian factor analysis approach to

combine the space-time disease mapping and joint modeling of different cancers. [9] used a full Bayesian hierarchical model to split the disease risks into shared and disease-specific spatiotemporal components. Their definition of multiple diseases was male and female subpopulations. [10] also applied a full Bayesian hierarchical approach to their spatiotemporal model to estimate the relative risk of various cancers while adjusting for age and gender. Using the hierarchical Bayesian factor models, [11] combined the dynamic factor analytic models with space-time disease mapping and produced a flexible framework for jointly analyzing multiple related diseases. Another study by [12] compared three formulations of the spatiotemporal shared component model on five diseases and examined the changes in the shared factors over time.

Incorporating the spatiotemporal modeling approaches to modeling of multiple diseases simultaneously adds significant complexity to the model structure [13]. The Bayesian hierarchical modeling approach makes an appropriate framework for solving complexities in the spatiotemporal structures [9,10,14,15]. The uniqueness of the Bayesian approach that differentiates it from other classical approaches is its robustness to interpretation of the posterior estimates and generating inferences of all model parameters [16].

The feasibility of using case notifications as a surrogate of population-based studies to estimate the shared and disease-specific risks of multiple diseases is also unknown. Against this background, we investigate the spatial and temporal patterns of Human Immunodeficiency Virus (HIV) and Tuberculosis (TB) burden in Kenya jointly using case notification data for a six-year period and characterize the areas with unusually high relative risks. These model maps describe new exposure hypotheses that warrant further epidemiological investigations on existing challenges and opportunities for disease surveillance and etiology

Motivation of the study

The HIV and TB diseases have a co-epidemic overlap in their epidemiologic characteristics and clinical manifestations [17]. Both the HIV and TB pathogens interact synergistically, accelerating the progress of illness thereby increasing the likelihood of death [18]. TB is a communicable disease and a major cause of ill health globally that affects the lungs (pulmonary) but can also affect other sites (extrapulmonary) [19,20]. Despite being a curable and preventable disease, TB is the leading cause of death from bacterial infection worldwide and undoubtedly representing a global public health priority [21–24]. TB is also the leading cause of morbidity and mortality among people living with HIV/AIDS [24], accounting for approximately 40% of deaths globally [20]. HIV disease is equally a major global health concern causing substantial morbidity, mortality rate, human suffering, and development challenges [25]. The rapid disease progression and associated leading opportunistic infection, TB, contribute to the high mortality rates [26]. Beyond affecting the health of individuals, HIV and TB diseases are stigmatized [17] and have a significant impact on the social, economic and political stability of the hardest-hit countries [27,28].

Kenya suffers from the dual epidemics ranking 4th and 15th globally in the high disease burden for HIV and TB respectively [29–31]. The high HIV prevalence in Kenya is the major driver of TB related mortality and TB prevalence is also the leading cause of HIV related mortality [31]. Studies on TB and HIV incidences in Kenya have limited spatial and temporal scoping making generalizations of their findings to the whole country difficult. This study uses routine case notification data to provides more accurate estimates and insights on the elevated risk areas of HIV and TB individually and jointly over time. By using the spatiotemporal approach to modeling HIV and TB diseases simultaneously, we present the shared and disease-specific

spatial risk patterns and explore their temporal evolution. The joint model also determines the combined and disease-specific elevated risk areas.

Materials and methods

Study location

The study was conducted in Kenya, a country of great diversity situated in East Africa extending between latitudes $4^0$30□N and $4^0$30□S and longitudes $34^0$00□E and $42^0$00□E [32]. Kenya has a coastline stretch of approximately 14420 km along the Indian ocean mostly covered by salt-tolerant mangrove trees that creates a distinctive ecological zone [33]. The total coverage area of Kenya is 583,367km^2 with 569,140km^2 as land area and 14,227km^2 being water area [34]. The diversity of Kenya's landscape is shaped by four distinguishable relief zones; these are the coastal and eastern broad plains, the central and western highlands, the Rift Valley and Lake Victoria basins [35].

Kenya is administratively subdivided into 47 counties as the first level of administrative subdivisions which in turn are further subdivided into 290 sub-counties and 1450 wards as the second and third levels of administrative subdivisions respectively [36]. The population of Kenya is unevenly distributed throughout the country and predominantly urban. The population density has remained on the increase from 77.9% in 2012 to 88.2% in 2017 per sq.km [37–39]. Supplementary information S2 presents the geospatial arrangements and the list of the 47 counties of Kenya according to their corresponding geographic codes as used in this study. The population estimates per county from 2012 – 2017 are in supplementary information S3.

A big obstacle to human capital development in Kenya is health challenges. Many people are exposed to a wide range of disease burdens largely because of the country's geographical, economic, political and climatic conditions [40] which are further compounded by inadequate resources to mitigate the impact of health risks [41]. Presently, cases of emerging and re-emerging diseases – like TB and HIV - are on the rise thereby having important implications on public health policy processes [42]. Understanding variations in incidence trends at the county level – where health services are planned, organized and delivered – is essential in addressing health inequalities.

Data Sources

This study considered routine case notification data on TB and HIV diseases in 47 counties of Kenya for 6 years, 2012-2017. Case notifications are data from specific subpopulations who seek treatment and care from health facilities; these are geographically representative of nearby populations. The data were collected and made available by two main sources; the National Tuberculosis Leprosy & Lung Disease Program (NTLD-P) and National AIDS & STIs Control Program (NASCOP). The Government of Kenya has routine case-based monitoring and reporting regulations for TB and HIV diseases through the NLTP and NASCOP programs. The Integrated Electronic Medical Records (EMR) Data Warehouse (IDWH) is the on-line case-based repository hosted by NASCOP accommodating all EMR databases from all health facilities across the country. It operates both as a repository and analytics platform presenting data through interactive dashboards and ad-hoc data analysis. Health facilities update their EMR databases into the IDWH on a monthly basis. The Tuberculosis Information from Basic Unit (TIBU) is the centrally located case-based surveillance system hosted by NLTP that allows for real-time reporting. Since its inception in 2012, TIBU has made notifications of TB patients very

timely and instant in report generation. All public, faith-based and private treatment centers in the country enter data into the TIBU system. There are 301 TB control zones across the 47 counties of Kenya which are coordinated by the Sub County TB and Leprosy Coordinators (SCTLCs), who are responsible for notifying TB cases from health facilities in their control zones into the TIBU system. Both programs have adapted the data recording and reporting standards of WHO at the health facilities in every county and the national surveillance system. For the purpose of this study, we analyze the data aggregated per county per year for the period 2012-2017.

Ethical considerations

This study involved the use of non-identifiable secondary data collected as part of the routine programs monitoring. Ethical permission to use the routinely collected data was obtained from NASCOP and NLTP, which are commissioned by the Ministry of Health to host the data surveillance systems for the HIV and TB programmes respectively. The study was subjected to Human Research Protection Review by the African Medical Research Foundation (AMREF Health Africa) Ethical and Scientific Review Committee (ESRC), which determined it not to constitute human participation.

Hierarchical Model specification

We formulated a statistical model that combined the spatiotemporal methods to the modeling of HIV and TB simultaneously. The study applied a full Bayesian Hierarchical approach on the model to estimate the spatial and temporal parameters for the two diseases individually and jointly. We applied the spatial and temporal specifications of [9] and [43] to our model. While [9] mapped a single disease for two subpopulations, we mapped two diseases -HIV and TB- with

co-epidemic overlap. They were also keen to interpret the unexpected differential risks between two subpopulations; our aim was to determine the shared and specific spatiotemporal patterns to interpret the relationships between the two diseases. Unlike [43], our model accounts for the joint space-time interaction using similar specifications as [9].

a. Log-linear model

The initial step in defining our model within the Hierarchical Bayesian framework was selecting the probability distribution for the observed data. This study utilized the Poisson distribution from the exponential family. For county s in the year t, we modeled cases notification y_{dst} for disease d, where d=1 is HIV and d=2 is TB as;

$$y_{dst} \sim Poisson(\mu_{dst}), \qquad \mu_{dst} = \rho_{dst} E_{dst}.$$

This study defined the mean μ_{dst} in terms of the unknown relative risk ρ_{dst} and the expected number of cases E_{dst}. We computed E_{dst} per county per year. Our statistical consideration for the standard population N was the average of the pooled county population estimates, i.e. $N = \frac{P}{T}$, where P is the estimated population of county s, which we are considering to be the population at risk of disease d and T is the number of years, which is six years for this study. We calculated the crude rate as $R_{dst} = \frac{\sum y_{dst}}{P_{dst}}$, where $\sum y_{dst}$ and P_{dst} are the number of cases for disease d and estimated population respectively for county s in the year t. We then multiplied the crude rate by the standard population to obtain the expected number of cases for disease d for county s in the year t as;

$$E_{dst} = R_{dst} \times N$$

The linear predictor of the unknown relative risk was on the logarithmic scale, $\eta_{dst} = \log(\rho_{dst})$ which is the recommended invertible link function for the Poisson family of distributions. The variability of the cases around the unknown relative risks ρ_{dst} for HIV and TB respectively were as follows;

$$y_{1st} \sim Pois(\rho_{1st} E_{1st}) \qquad \eta_{1st} = \alpha_1 + \lambda_s \delta + \xi_t \kappa + \beta_{1s} + \gamma_{1t} + \upsilon_{st}$$

$$y_{2st} \sim Pois(\rho_{2st} E_{2st}) \qquad \eta_{2st} = \alpha_2 + \frac{\lambda_s}{\delta} + \frac{\xi_t}{\kappa} + \beta_{2s} + \gamma_{2t} + \upsilon_{st}$$

The linear predictor was defined by the following terms; the shared spatial effect, ($\boldsymbol{\lambda} = \{\lambda_s\}_{s=1,2,\ldots,S}$), the disease-specific spatial effect ($\boldsymbol{\beta} = \{\beta_{ds}\}_{d=1,2;\ s=1,2,\ldots,S}$), the shared time trend ($\boldsymbol{\xi} = \{\xi_t\}_{t=1,\ldots,T}$), the disease-specific time trend ($\boldsymbol{\gamma} = \{\gamma_{dt}\}_{d=1,2;\ t=1,\ldots,T}$), and the space-time interaction term ($\boldsymbol{\upsilon} = \{\upsilon_{st}\}_{s=1,\ldots,S;\ t=1,\ldots,T}$). The notations α_d, λ_s, and ξ_t captured disease-specific intercept, space and time main effects respectively whereas γ_{dt} and β_{ds} were disease-time and disease-space interactions of order 2 respectively. The coefficients δ and κ represented the spatial and temporal scaling parameters on the shared term to the risk of TB compared to HIV. Even though the overall relative risk level is the same for both diseases, the magnitude of the area-specific and time-specific relative risks may differ - hence the need for the scaling parameters [9,44,45]. Therefore, contribution of the shared component to the overall relative risk is weighted by the scaling parameters to allow different risk gradients for each disease.

We applied a symmetric formulation to both the shared and disease-specific random effects, implying that λ_s, and ξ_t captured the common spatial and temporal patterns. The terms γ_{dt} and β_{ds} allowed for departures from the shared patterns for the different diseases. The space-time interaction term, υ_{st} provided additional flexibility towards identifying varying patterns. We

assumed the shared random effects to be the structured effects and the disease-specific random effects to be unstructured effects

b. Bayesian prior specification

In a Bayesian framework, random effects are unknown quantities assigned to prior distributions that reflect any prior knowledge on the structure of the effects. The model assigned priors and generated the posterior distribution used for deriving the conditional densities for posterior sampling.

The spatial random effects λ_s and β_s assumed a spatially correlated prior distribution (CAR spatial priors) with a neighborhood matrix W defined by contiguity; τ_λ and τ_β were the prior hyperparameters. The CAR spatial prior defines a binary specification for the geographical contiguity such that correlation is certain for geographically adjacent areas that is:

$$S_{kj} = \begin{cases} 1 & if\ k \sim j \\ 0 & otherwise \end{cases}$$

The non-contiguous areal units are conditionally independent given the values of the remaining random effects. To reflect the prior of yearly fluctuations for ξ_t and γ_{dt}, the study assumed a random walk prior of order 1 (RW (1)) with a weighted matrix Q which defines the temporal neighborhood with τ_ξ and τ_γ as the prior hyperparameters. The space-time interaction term ν_{st} prior was a simple exchangeable hierarchical structure $\nu_{st} \sim N(0, \tau_\nu)$. We defined improper priors for the intercept as $\alpha \sim N(0,0)$. The scaling parameters assume normal priors $N(0, \sigma^2)$ which are symmetric around zero, therefore, any value is as equally likely as the reciprocal value and the posterior distribution of the relative risks for each disease are exactly the same. For the distribution of the hyper-parameters, we assumed the default specifications of INLA whereby we

assigned minimally informative priors on the log of the precision of both the structured and unstructured effects, logGamma(1, 0.0005). INLA estimates the posterior marginal distribution for the hyperparameters using an integration-free algorithm described in [46]. We use INLA approach for the model estimations as it is capable of handling complex models with large predictor spaces. Equally the approach does not require convergence checking (unlike McMC) [47,48] as it does not suffer from slow convergence and poor mixing.

Results

To determine the areas of high risks, we created the spatial maps of the standardized incidence ratio for HIV (Fig 1) and TB (Fig 2). The two diseases displayed varying spatiotemporal patterns, though most of the regions of high risk for HIV were also high risk for TB. The progression of the risk during the period 2012 – 2017 was much faster in TB as compared to HIV

Fig 1. Standardized incidence ratio $\left(\frac{O_{dst}}{E_{dst}}\right)$ for HIV

Fig 2. Standardized incidence ratio $\left(\frac{O_{dst}}{E_{dst}}\right)$for TB

We considered the analysis of the combined spatial patterns in the model. Table 1 provides the summary statistics of the shared and disease-specific spatial effects while Fig 3 presents the relative risk of the combined spatial patterns for HIV and TB. The estimates of the different spatial effects are statistically significant at 5% as depicted by the credible intervals. Both HIV and TB had similar areas of high risk in the South West, Central and South regions of Kenya. Further, TB revealed additional areas of high risk in the North West and North of Kenya.

Table 1. Summary statistics of the shared and disease-specific spatial and temporal effects

Parameters	mean	2.5%	97.5%
Fixed effects:			
α_1	0.660	0.622	0.700
α_2	0.885	0.872	0.899

Random effects:			
$\tau_{\beta_{1s}}$	2.301	1.566	4.191
$\tau_{\beta_{2s}}$	24.557	4.389	588.749
τ_{λ_s}	1.319	1.121	1.738
$\tau_{\gamma_{1t}}$	3.222	2.611	139.491
$\tau_{\gamma_{2t}}$	53.038	2.464	831.308
τ_{ξ_t}	1.030	1.003	1.138
$\tau_{\upsilon_{st}}$	3.219	2.561	4.568

Fig 3. Relative risk of the combined spatial patterns ($\lambda_s + \beta_{ds}$)

The posterior means of the shared and disease-specific spatial patterns are presented in Fig 4. In all the maps, common areas of high risk were in the west part of the country. Equally, the shared spatial pattern showed fewer regions of high risk compared to TB-specific spatial patterns. This could be because of higher dependence of HIV on the shared spatial term making the shared pattern to account for most of HIV spatial patterns

Fig 4. Relative risk of shared spatial patterns (λ_s) and disease specific spatial patterns (β_{ds})

The posterior means of the shared, disease-specific, and combined temporal trends are in Fig 5. The shared temporal effect displayed the overall decreasing risk trend in time with estimates close to zero for most years. The HIV temporal trend exhibited increasing risk over time displaying relative risks between 0.8 - 1.2 whereas the TB temporal trend presented a nearly constant trend across the years with relative risks close to one. The combined temporal trends for HIV and TB showed an increasing risk trend for both diseases. From the combined temporal trends graph, HIV risk was relatively lower than TB in 2012 and 2013. From 2015 - 2017, the HIV risk trend surpassed that of TB. Even though the shared effect shows minimal evidence on the joint diseases, the disease-specific and combined temporal effects show very similar temporal pattern.

Fig 5. Relative risk of the shared ξ_t, specific γ_{dt} and the combined temporal trends ($\xi_t + \gamma_{dt}$)

The posterior probabilities for the smoothed joint Spatio-temporal interaction $P\left(e^{(\upsilon_{st})} > 1|y\right)$ are in Fig 6, and the posterior estimates are presented in Table 1. Over the years the joint incidence risk reduced progressively with all spatial regions having a relative risk less than 1.

Fig 6. Posterior probabilities for the joint Spatio-temporal interaction $P\left(e^{(\upsilon_{st})} > 1|y\right)$

Fig 7(a) displays the spatial pattern of the posterior mean for the joint disease risk $\zeta_s = e^{(\lambda_s + \beta_{ds})}$. The uncertainty associated with the posterior means, which lays in the exceedance risk ζ_s: P (ζ_s>1|y), is presented in Fig 7(b). The extreme West and Central regions depicted elevated joint relative risks characterized by a spatial relative risk above one and exceedance probability above 0.8

Fig 7. Maps of the spatial pattern of joint disease risk $\left(\zeta_s = e^{(\lambda_s + \beta_{ds})}\right)$ and uncertainty for the spatial effect $\left(\zeta_s : P(\zeta_s > 1|y)\right)$

Discussion

We presented a Bayesian Hierarchical approach to the joint modeling of spatio-temporal routinely collected health data. The joint modeling approaches have yielded substantial co-dynamic insights via mathematical, statistical and computational approaches [49]. By optimizing the spatial scale at different points in time, spatial heterogeneity influences the interpretation of temporal patterns more especially in disease dynamics and surveillance [49]. This is especially true for the case of HIV and TB that have significant geographic overlap and are subject to diverse regional variations in their co-dynamics. HIV and TB rank as the leading causes of death from infectious diseases globally with an estimated 2.5 million new HIV infections and 8.7 million incidences of TB annually [50,51]. They have a close link even though their biological co-existence and co-dynamics vary regionally with much burden in Sub-Saharan Africa [31,52]. This study determined the space-time joint risk trends of HIV and TB in Kenya. Our model

295 enabled us to define the shared and specific spatial and temporal patterns of HIV and TB thus
296 identifying similarities and differences in the distribution of the relative risks associated with
297 each disease. The model separately estimated the shared and disease-specific relative risks and
298 displayed the spatial-disease, temporal-disease, and spatio-temporal disease interaction effects
299 across all regions. We included scaling components on the shared spatial and temporal
300 parameters to compare their strength signals for HIV and TB.

301 The disease-specific spatial and temporal patterns detected areas with varying spatial trends and
302 temporal variations for each disease. The HIV high-risk areas were to the further west of Kenya
303 spreading towards the central and further south. The TB high-risk areas were similar to the HIV
304 high-risk areas but also spread upwards towards the North. The TB geographical progression in
305 relation to HIV was proportionally higher which could reflect environmental factors favoring the
306 TB spread in the high-density settlements especially in the towards the North. These findings are
307 corroborated in other studies by [31,53]. Looking beyond Kenya, studies by [54] revealed that
308 TB appeared to outpace HIV in Rwanda and Burundi while HIV greatly outpaced TB in
309 Mauritania, Senegal and The Gambia

310 Joint temporal analysis is important when investigating the temporal coherency of
311 epidemiological trends from the same area [54]. In our study, the shared temporal trend
312 presented a steep decrease from 2012 to 2013 then an almost constant risk without any
313 significant variation over time. However, the disease-specific and combined temporal trends
314 presented an increased risk over time. The temporal trend of HIV risk was lower than that of TB
315 for the years 2012 and 2013 but between 2015 and 2017 the HIV risk was higher than TB risk.
316 Similar studies in Sub-Saharan Africa that utilized routinely collected data observed similar
317 temporal dynamics [55–58]. A possible explanation could be HIV drives TB related incidences,

318 therefore, the incidence and prevalence of TB increases (decreases) with increasing (decreasing)
319 HIV trends [59–61].

320 Our study successfully detected the spatial congruence in the distribution of TB and HIV in
321 approximately 29 counties around the western, central and southern regions of Kenya. The
322 spatial patterns were largely similar for Homabay, Siaya, Kisumu, Busia and Migori counties as
323 the high risk with Mandera, Wajir and Garissa counties at low risk for both HIV and TB. The
324 distribution of the shared relative risks had minimal difference with the HIV disease-specific
325 relative risk whereas that of TB presented many more counties as high-risk areas. This could be
326 attributed to higher dependence of HIV on the shared spatial term making the shared pattern
327 account for most HIV spatial patterns. Similar studies by [62] in China and [53] in Uganda
328 observed significantly persistent clusters for TB and HIV using the spatial co-clustering
329 approach. They examined the clusters exhibited by each disease as well as the combined.

330 Our model also presented areas of elevated joint risk by looking at the posterior probabilities of
331 the relative risk to detect the joint geographical patterns. There was strong evidence of disease-
332 time, disease-space and space-time interactions. Studies by [62] and [63] also found a strong
333 association on the joint risks of HIV and TB they used bivariate maps to show that the joint
334 distribution for both TB and HIV diseases was spatially heterogeneous across Brazil.

335 The primary limitation of the study is using routine case notification data as a surrogate measure
336 of the general population of infected persons; this type of data has challenges of underreporting.
337 The spatial regions were based on the counties since there was no data for sub-county or health
338 facility levels. The age and sex standardized rates were not utilized as the HIV data was not
339 disaggregated by age and gender. Whereas this kind of data is not completely spatially random

for the joint epidemic burden, it still captures the spatiotemporal patterns of incidence risk, which is the ultimate goal of this study.

Conclusions

We defined a joint Bayesian space-time model of two related diseases to jointly quantify the risk of TB relative to HIV thereby facilitating the comparative benefits obtained across populations. The disease burden was apparent at each spatial level of analysis. Identifying the spatial and temporal similarities between HIV and TB enabled us to understand the shared risk. The flexibility and informative outputs of Bayesian Hierarchical Models played a key role in clustering these risk areas. Quantifying how HIV and TB varied together provided additional insights towards collaborative monitoring of the diseases and control efforts. To control the HIV-TB twin epidemic it is important to determine the TB life history that could greatly reduce TB incidences if intervened and which stages of HIV needs optimized TB control benefits

Acknowledgments

The authors sincerely acknowledge the National Tuberculosis, Leprosy and Lung Cancer Program (NLTP) and the National AIDS and STI Control Program (NASCOP) for their cordial support in providing the data used in this study

References

1. Riebler A, Sørbye SH, Simpson D, Rue H. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Stat Methods Med Res. 2016;25(4):1145–65.

2. Schrödle B, Held L. A Primer on Disease Mapping and Ecological Regression Using

INLA. Comput Stat. 2010;26(2):241–58.

3. Ugarte M, Adin A, Goicoa T, Militino A. On fitting spatio-temporal disease mapping models using approximate Bayesian inference. Stat Methods Med Res. 2014;23(6):507–30.

4. Dabney AR, Wakefield JC. Issues in the mapping of two diseases. Stat Methods Med Res. 2005;14(1):83–112.

5. Baker J, White N, Mengersen K, Rolfe M, Morgan GG. Joint modelling of potentially avoidable hospitalisation for five diseases accounting for spatiotemporal effects: A case study in New South Wales, Australia. PLoS One. 2017;12(8):e0183653.

6. Rotejanaprasert C, Lawson AB, Iamsirithaworn S. Spatiotemporal multi-disease transmission dynamic measure for emerging diseases: an application to dengue and zika integrated surveillance in Thailand. BMC Med Res Methodol. 2019;19(1):200.

7. Earnest A, Beard JR, Morgan G, Lincoln D, Summerhayes R, Donoghue D, et al. Small area estimation of sparse disease counts using shared component models-application to birth defect registry data in New South Wales, Australia. Heal Place. 2010;16(4):684–93.

8. Tzala E. Multivariate analysis of spatial and temporal variation in cancer mortality. Imperial College London; 2004.

9. Richardson S, Abellan JJ, Best N. Bayesian spatio-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (UK). Stat Methods Med Res. 2006;15(4):385–407.

10. Oleson JJ, Smith BJ, Kim H. Joint Spatio-Temporal Modeling of Low Incidence Cancers

Sharing Common Risk Factors. J Data Sci. 2008;6:105–23.

11. Tzala E, Best N. Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer mortality. Stat Methods Med Res. 2008;17(1):97–118.

12. Baker JF. Bayesian Spatiotemporal Modelling of Chronic Disease Outcomes. Queensland University of Technology; 2017.

13. Meliker JR, Sloan CD. Spatio-temporal epidemiology: Principles and opportunities. Spat Spatiotemporal Epidemiol. 2011;2(1):1–9.

14. Lee D. A comparison of conditional autoregressive models used in Bayesian disease mapping. Spat Spatiotemporal Epidemiol. 2011;2:79–89.

15. Bivand RS, Pebesma E, Gomes-Rubio V. Applied Spatial Data Analysis with R. 2nd ed. Gentleman R, Hornik K, Parmigiani GG, editors. Springer; 2013.

16. Musenge E. Modeling spatiotemporal patterns of childhood HIV/TB related mortality and malnutrition: Applications to agincourt data in rural South Africa. University of the Witwatersrand. University of the Witwatersrand; 2013.

17. Tsiouris SJ, Gandhi NR, El-Sadr WM, Gerald F. Tuberculosis and HIV-Needed: A New Paradigm for the Control and Management of Linked Epidemics. J Int AIDS Soc. 2007;9(3):62.

18. Mayer KH, Hamilton CD. Synergistic Pandemics: Confronting the Global HIV and Tuberculosis Epidemics. Clin Infect Dis. 2010;50(s3):S67–70.

19. WHO. Global tuberculosis report 2018. Geneva (Switzerland); 2018.

20. Macneil A, Glaziou P, Sismanidis C, Maloney S, Floyd K. Global epidemiology of

tuberculosis and progress toward achieving global targets — 2017. Morb Mortal Wkly Rep. 2019;68(11):263–6.

21. Garcia-saenz A, Saez M, Napp S, Casal J, Luis J, Acevedo P, et al. Spatio-temporal variability of bovine tuberculosis eradication in Spain (2006–2011). Spat Spatiotemporal Epidemiol. 2014;10:1–10.

22. Jacobs AJ, Mongkolsapaya J, Screaton GR, McShane H, Wilkinson RJ. Antibodies and tuberculosis. Tuberculosis. 2016;101:102–13.

23. Saunders MJ, Tovar MA, Datta S, Evans BEW, Wingfield T, Evans CA. Pragmatic tuberculosis prevention policies for primary care in low- and middle-income countries. Eur Respir J. 2018;51(3):1800315.

24. UNAIDS. Joint United Nations Programme on HIV/AIDS. Global report: UNAIDS report on the global AIDS epidemic 2010. 2010.

25. Mgori NK, Mash R. HIV and/or AIDS-related deaths and modifiable risk factors: A descriptive study of medical admissions at Oshakati Intermediate Hospital in Northern Namibia. African J Prim Heal Care Fam Med. 2015;7(1):e1–7.

26. Burrage A, Patel M, Mirkovic K, Dziuban E, Teferi W, Broyles L, et al. Trends in Antiretroviral Therapy Eligibility and Coverage Among Children Aged <15 Years with HIV Infection - 20 PEPFAR-Supported Sub-Saharan African Countries, 2012-2016. MMWR Morb Mortal Wkly Rep. 2018;67(19):552–5.

27. HIV.gov. Global Statistics [Internet]. 2018. Available from: https://www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics

423 28. KFF. Global HIV/AIDS Epidemic [Internet]. 2018. Available from:
424 https://www.kff.org/global-health-policy/fact-sheet/the-global-hivaids-epidemic/

425 29. Otiende V, Achia T, Mwambi H. Bayesian modeling of spatiotemporal patterns of TB-
426 HIV co-infection risk in Kenya. BMC Infect Dis. 2019;19:902.

427 30. Enos M, Sitienei J, Ong'ang'o J, Mungai B, Kamene M, Wambugu J, et al. Kenya
428 tuberculosis prevalence survey 2016: Challenges and opportunities of ending TB in
429 Kenya. Hozbor DF, editor. PLoS One. 2018;13(12):e0209098.

430 31. Onyango DO, Yuen CM, Cain KP, Ngari F, Masini EO, Borgdorff MW. Reduction of
431 HIV-associated excess mortality by antiretroviral treatment among tuberculosis patients in
432 Kenya. Cameron DW, editor. PLoS One. 2017;12(11):e0188235.

433 32. Google Earth. Map showing location of Kenya [Internet]. Available from:
434 https://www.google.com/earth/

435 33. Yap W, Landoy R, Fisheries Department FAO. Report on a Survey of the Coastal Areas
436 of Kenya for Shrimp Farm Development [Internet]. 1986. Available from:
437 http://www.fao.org/3/AC574E00.htm

438 34. United Nations Office Nairobi. Introductory guide to Kenya [Internet]. Division of
439 Conference Services. Available from: https://dcs.unon.org/node/75

440 35. Mathu EM, Davies TC. Geology and the environment in Kenya. J African Earth Sci.
441 1996;23(4):511–39.

442 36. CoG-KE. The 47 Counties of Kenya [Internet]. Available from: http://cog.go.ke/the-47-
443 counties

444 37. Knoema. Kenya Population density, 1960-2018 [Internet]. Available from:
445 https://knoema.com/atlas/Kenya/Population-density

446 38. GoK, UNFPA - Kenya Country Office. Kenya Population Situation Analysis. 2013.

447 39. The World Bank. World Development Indicators (WDI) Data Catalog [Internet].
448 Available from: https://datacatalog.worldbank.org/dataset/world-development-indicators

449 40. Ministry of medical services, Ministry of public health & sanitation. The Kenya health
450 sector strategic & investment plan July 2012 - June 2017 [Internet]. 2017. Available from:
451 http://www.health.go.ke/wp-content/uploads/2016/03/MINISTERIAL-STRATEGIC-
452 INVESTMENT-PLAN.pdf

453 41. Institute for Health Metrics and Evaluation. Kenya Profile. Seattle, WA: IHME,
454 University of Washington [Internet]. 2018. Available from:
455 http://www.healthdata.org/Kenya

456 42. Achoki T, Miller-Petrie MK, Glenn SD, Kalra N, Lesego A, Gathecha GK, et al. Health
457 disparities across the counties of Kenya and implications for policy makers, 1990–2016: a
458 systematic analysis for the Global Burden of Disease Study 2016. Lancet Glob Heal.
459 2019;7(1):e81–95.

460 43. Gómez-Rubio V, Palmí-Perales F, López-Abente G, Ramis-Prieto R, Fernández-Navarro
461 P. Bayesian joint spatio-temporal analysis of multiple diseases. Stat Oper Res Trans.
462 2019;43(1):51–74.

463 44. Knorr-Held L, Best NG. A shared component model for detecting joint and selective
464 clustering of two diseases. J R Stat Soc Ser A (Statistics Soc. 2001;164(1):73–85.

45. Held L, Natár'io I, Fenton SE, Rue H, Becker N. Towards joint disease mapping. Stat Methods Med Res. 2005;14:61–82.

46. Martins TG, Simpson D, Lindgren F, Rue H. Bayesian computing with INLA: New features. Comput Stat Data Anal. 2013;67:68–83.

47. Günhan BK, Friede T, Held L. A design-by-treatment interaction model for network meta-analysis and meta-regression with integrated nested Laplace approximations. Res Synth Methods. 2018;9(2):179–94.

48. Gómez-Rubio V, Palmí-Perales F. Multivariate posterior inference for spatial models with the integrated nested Laplace approximation. J R Stat Soc Ser C (Applied Stat. 2019;68(1):199–215.

49. Sánchez MS, Lloyd-Smith JO, Williams BG, Porco TC, Ryan SJ, Borgdorff MW, et al. Incongruent HIV and tuberculosis co-dynamics in Kenya: interacting epidemics monitor each other. Epidemics. 2009;1(1):14–20.

50. WHO. A guide to monitoring and evaluation for collaborative TB/HIV activities. 2015.

51. Venturini E, Turkova A, Chiappini E, Galli L, de Martino M, Thorne C. Tuberculosis and HIV co-infection in children. BMC Infect Dis. 2014;14(Suppl 1):S5.

52. Gwitira I, Murwira A, Mberikunashe J, Masocha M. Spatial overlaps in the distribution of HIV/AIDS and malaria in Zimbabwe. BMC Infect Dis. 2018;18:598–607.

53. Aturinde A, Farnaghi M, Pilesjö P, Mansourian A. Spatial analysis of HIV-TB co-clustering in Uganda. BMC Infect Dis. 2019;19(1):612.

54. Sánchez MS, Lloyd-Smith JO, Getz WM. Monitoring linked epidemics: the case of

tuberculosis and HIV. PLoS One. 2010;5(1):e8796.

55. Assebe LF, Reda HL, Wubeneh AD, Lerebo WT, Lambert SM. The effect of isoniazid preventive therapy on incidence of tuberculosis among HIV-infected clients under pre-ART care, Jimma, Ethiopia: a retrospective cohort study. BMC Public Health. 2015;15:346–54.

56. Nakiyingi L, Ssengooba W, Nakanjako D, Armstrong D, Holshouser M, Kirenga BJ, et al. Predictors and outcomes of mycobacteremia among HIV-infected smear- negative presumptive tuberculosis patients in Uganda. BMC Infect Dis. 2015;15:62–9.

57. Sabasaba A, Mwambi H, Somi G, Ramadhani A, Mahande MJ. Effect of isoniazid preventive therapy on tuberculosis incidence and associated risk factors among HIV infected adults in Tanzania: a retrospective cohort study. BMC Infect Dis. 2019;19(1):62.

58. Yirdaw KD, Jerene D, Gashu Z, Edginton ME, Kumar AM V, Letamo Y, et al. Beneficial effect of isoniazid preventive therapy and antiretroviral therapy on the incidence of tuberculosis in people living with HIV in Ethiopia. PLoS One. 2014;9(8):e104557.

59. Geldmacher C, Zumla A, Hoelscher M. Interaction between HIV and Mycobacterium tuberculosis: HIV-1-induced CD4 T-cell depletion and the development of active tuberculosis. Curr Opin HIV AIDS. 2012;7(3):268–75.

60. Naing C, Mak JW, Maung M, Wong SF, Kassim AIBM. Meta-analysis: the association between HIV infection and extrapulmonary tuberculosis. Lung. 2013;191(1):27–34.

61. Pontillo A, Carvalho MS, Kamada AJ, Moura R, Schindler HC, Duarte AJS, et al. Susceptibility to Mycobacterium tuberculosis infection in HIV-positive patients is

associated with CARD8 genetic variant. J Acquir Immune Defic Syndr. 2013;63(2):147–51.

62. Wei W, Wei-Sheng Z, Ahan A, Ci Y, Wei-Wen Z, Ming-Qin C. The Characteristics of TB Epidemic and TB/HIV Co-Infection Epidemic: A 2007-2013 Retrospective Study in Urumqi, Xinjiang Province, China. PLoS One. 2016;11(10):e0164947.

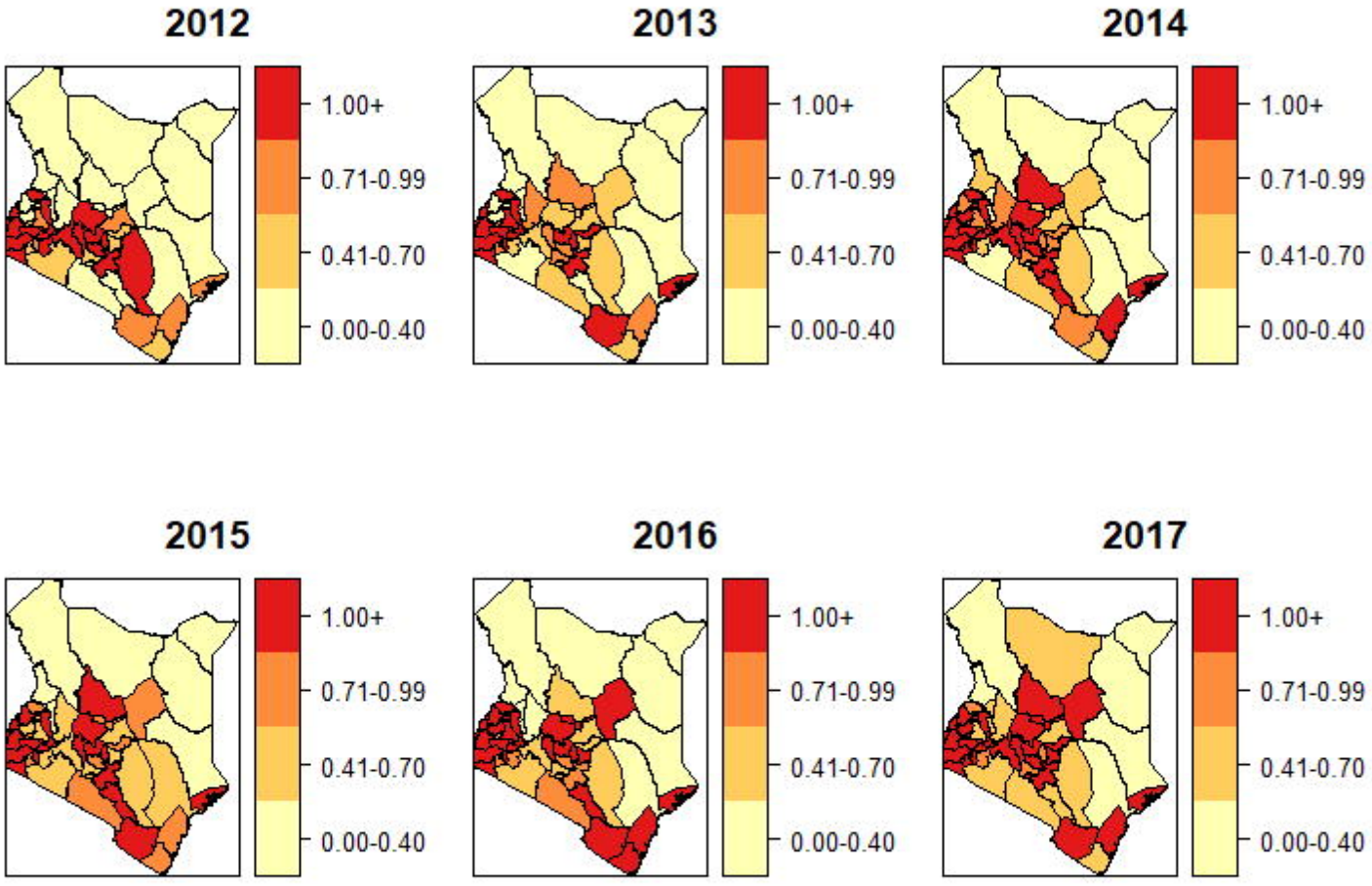
63. Ross JM, Henry NJ, Dwyer-Lindgren LA, de Paula Lobo A, Marinho de Souza F, Biehl MH, et al. Progress toward eliminating TB and HIV deaths in Brazil, 2001-2015: a spatial assessment. BMC Med. 2018;16:144–53.

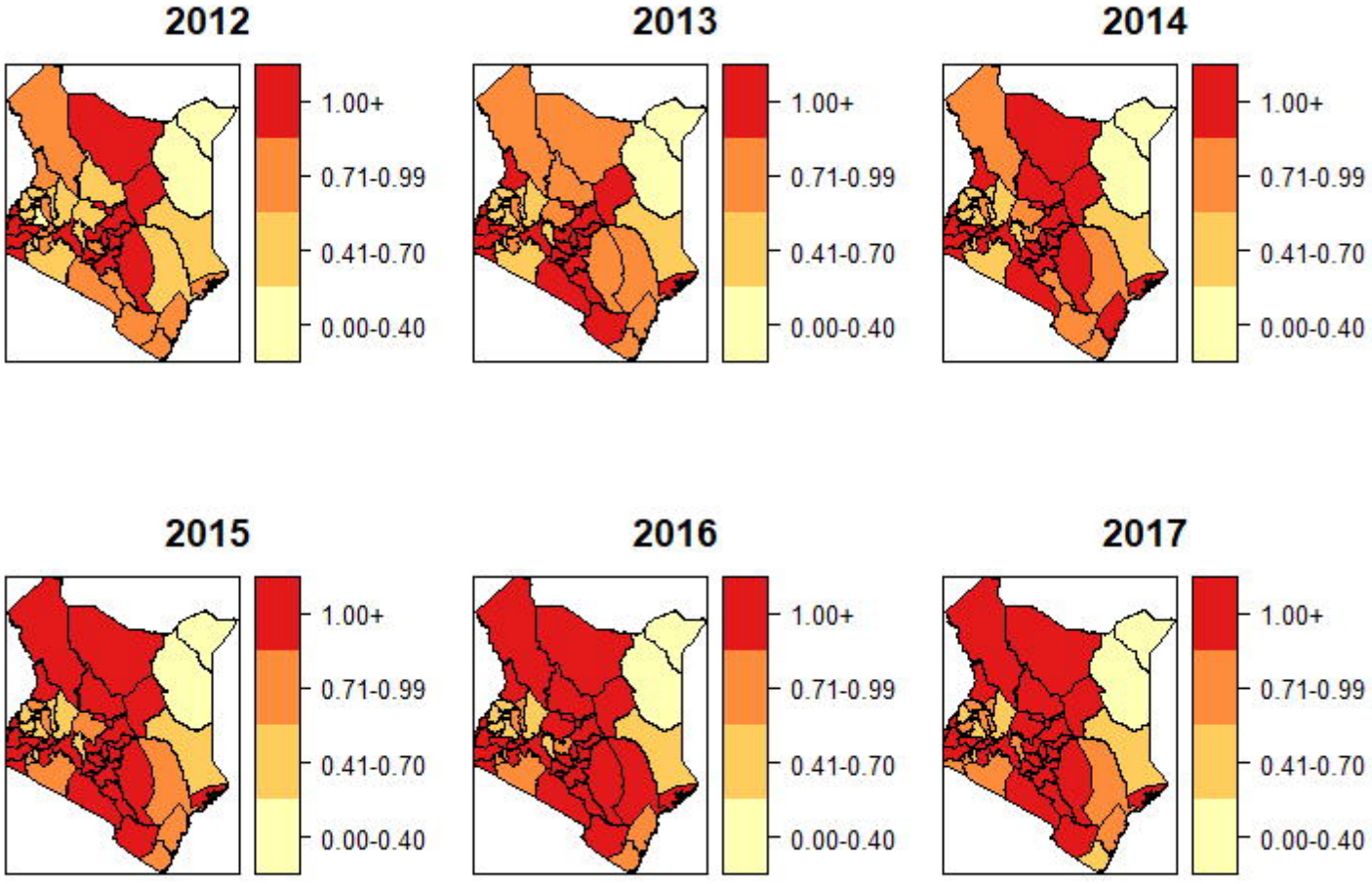
Supporting Information

S1 Text. R codes used in the analysis

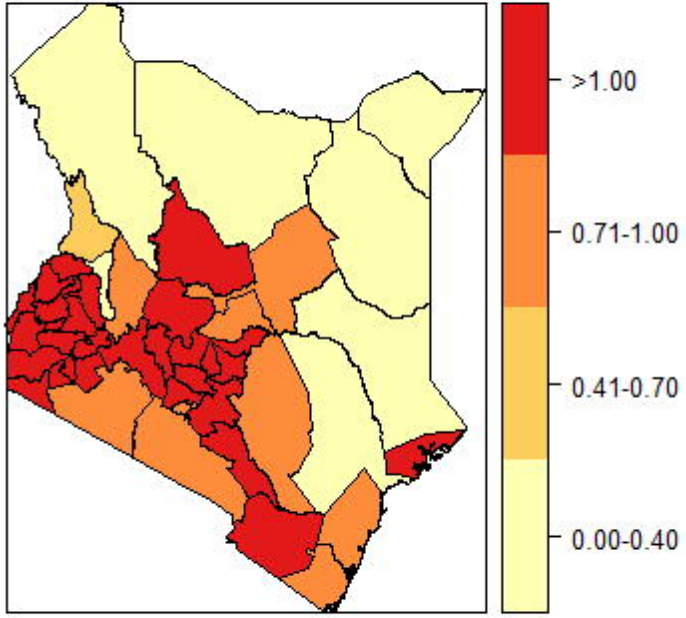
S2 Text. List of counties in Kenya

S3 Text. County population estimates 2012-2017

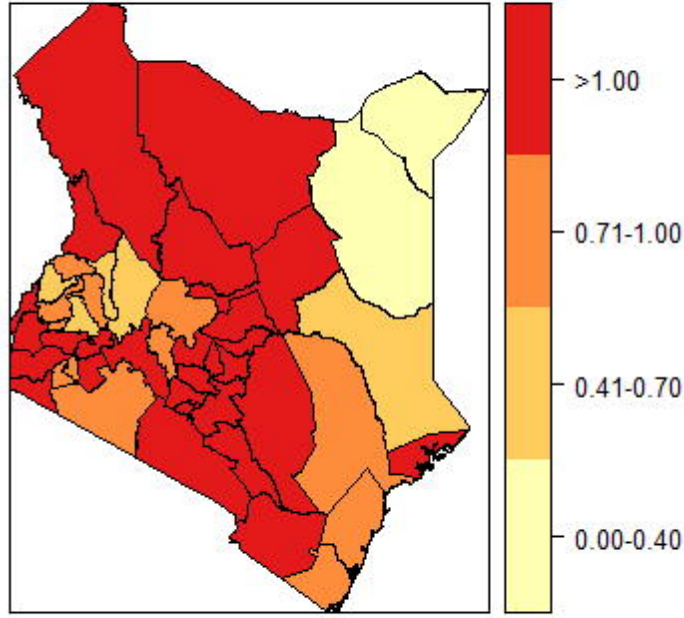

2012
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2013
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2014
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2015
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2016
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2017
1.00+
0.71-0.99
0.41-0.70
0.00-0.40

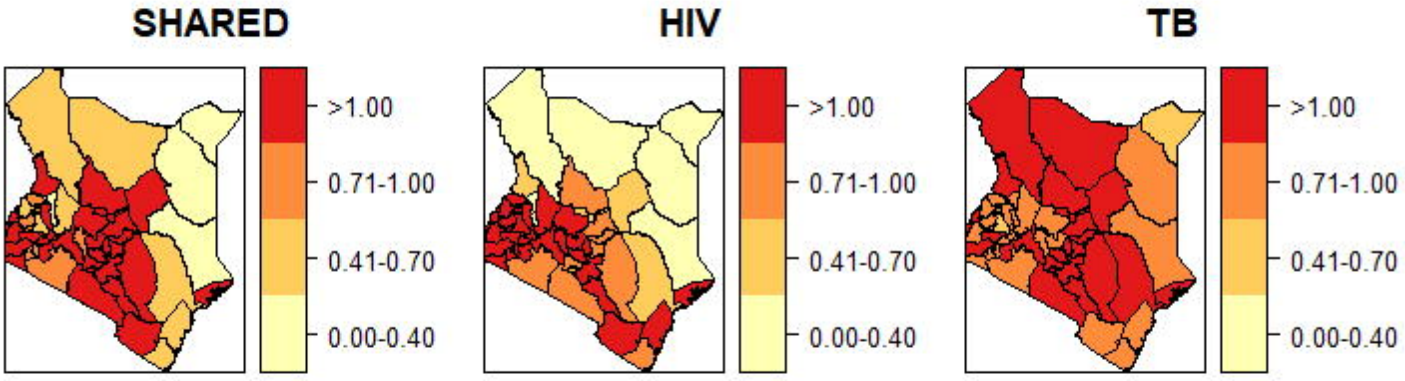

2012
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2013
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2014
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2015
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2016
1.00+
0.71-0.99
0.41-0.70
0.00-0.40
2017
1.00+
0.71-0.99
0.41-0.70
0.00-0.40

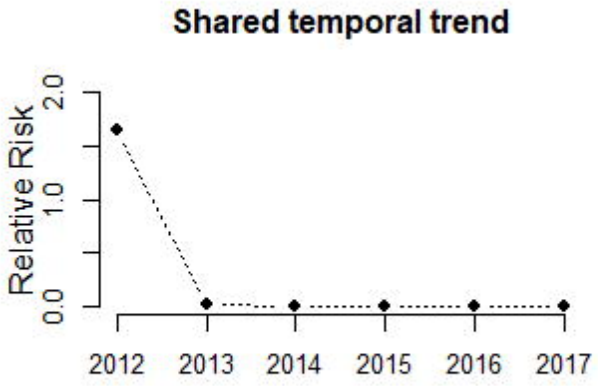
HIV Combined spatial patterns

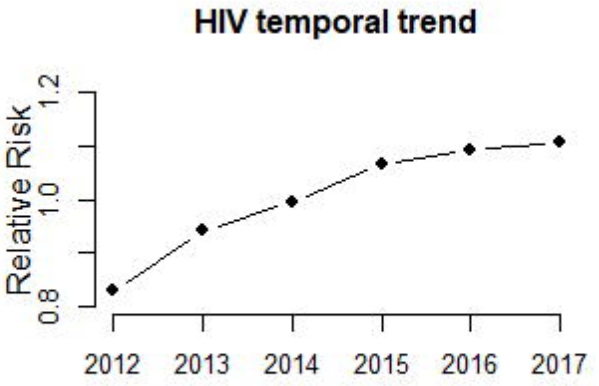


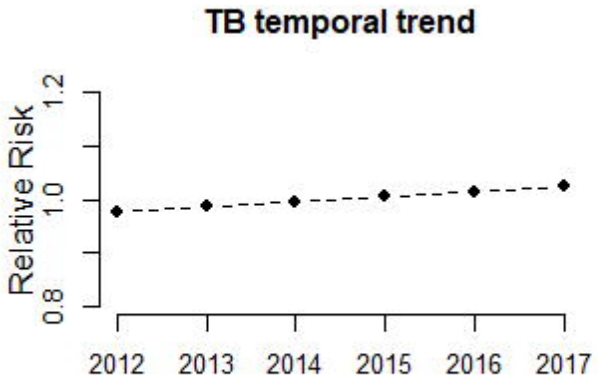
TB combined spatial patterns

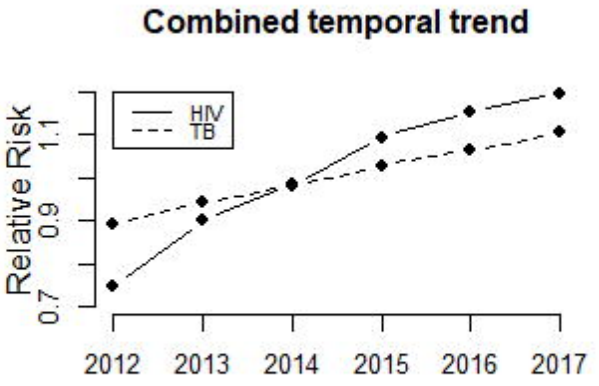


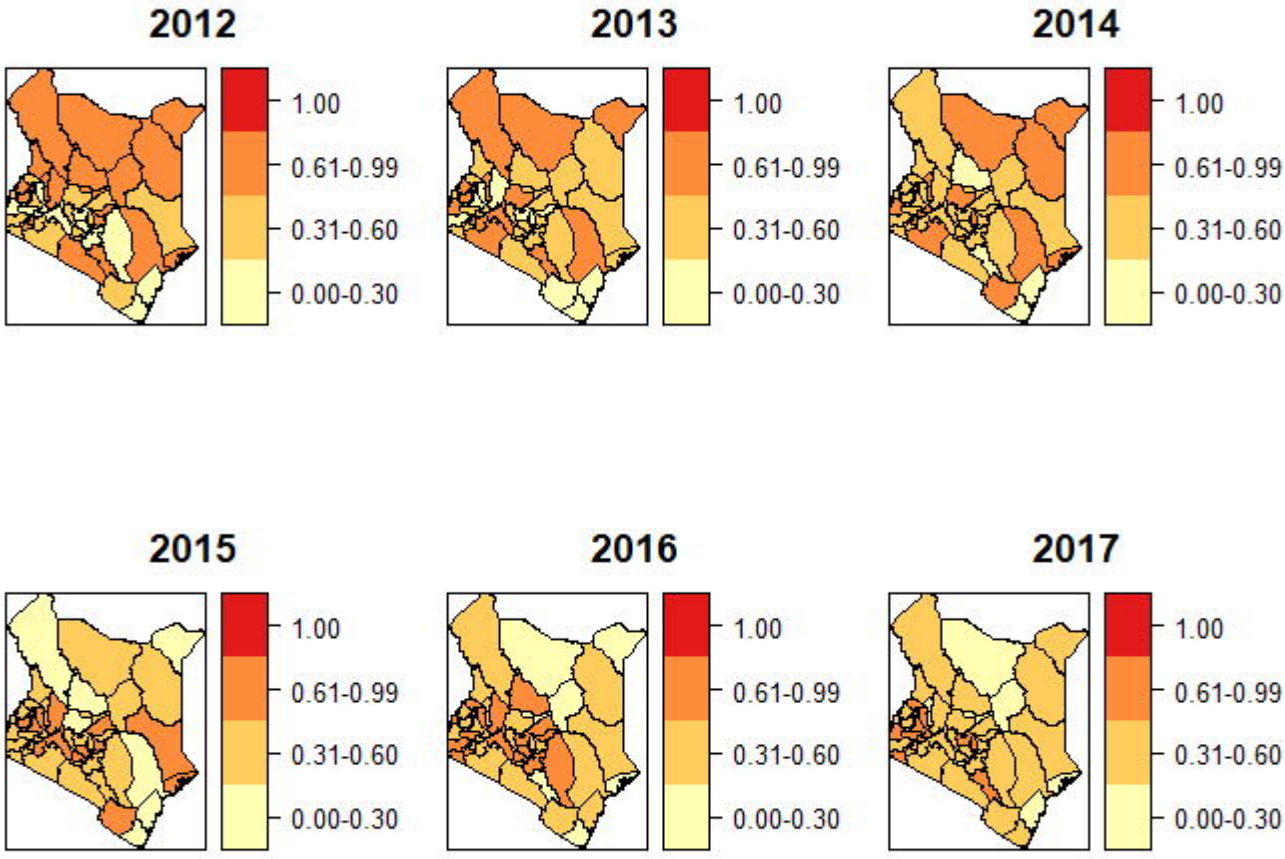

SHARED
>1.00
0.71-1.00
0.41-0.70
0.00-0.40
HIV
>1.00
0.71-1.00
0.41-0.70
0.00-0.40
TB
>1.00
0.71-1.00
0.41-0.70
0.00-0.40

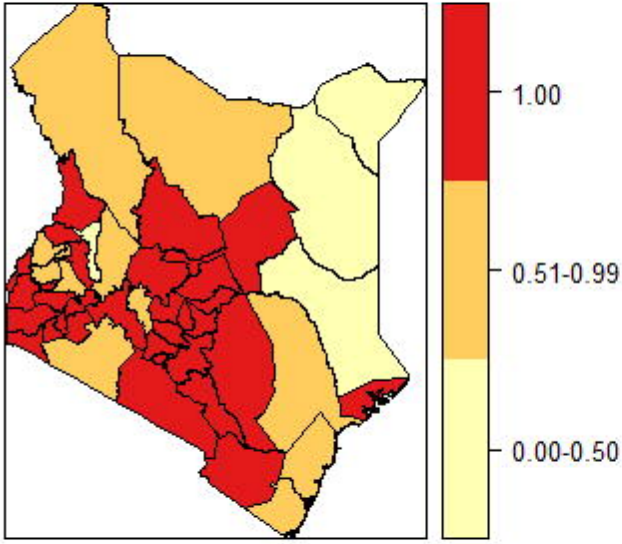

Shared temporal trend
Relative Risk
0.0
1.0
2.0
2012
2013
2014
2015
2016
2017

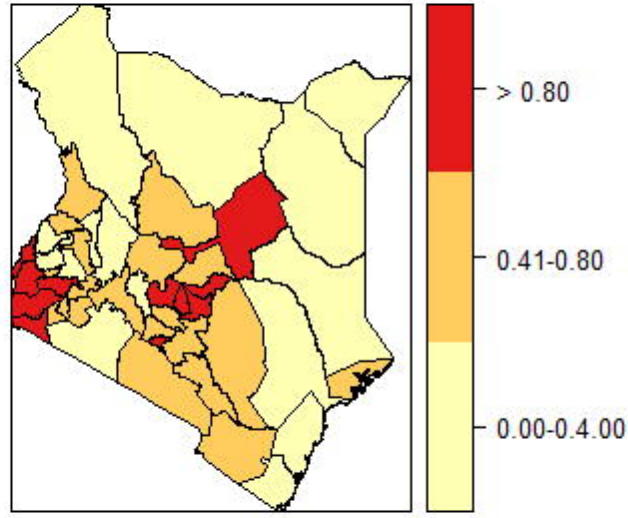

HIV temporal trend
Relative Risk
0.8
1.0
1.2
2012
2013
2014
2015
2016
2017


TB temporal trend
Relative Risk
0.8
1.0
1.2
2012
2013
2014
2015
2016
2017


Combined temporal trend
HIV
TB
Relative Risk
0.7
0.9
1.1
2012
2013
2014
2015
2016
2017


2012
1.00
0.61-0.99
0.31-0.60
0.00-0.30
2013
1.00
0.61-0.99
0.31-0.60
0.00-0.30
2014
1.00
0.61-0.99
0.31-0.60
0.00-0.30
2015
1.00
0.61-0.99
0.31-0.60
0.00-0.30
2016
1.00
0.61-0.99
0.31-0.60
0.00-0.30
2017
1.00
0.61-0.99
0.31-0.60
0.00-0.30


1.00
0.51-0.99
0.00-0.50


> 0.80
0.41-0.80
0.00-0.4.00