Estimating the spatial distribution of acute undifferentiated fever (AUF) and associated risk factors using emergency…

6 authors, including:

Boris Kauhl
AOK Nordost - Die Gesundheitskasse

11 PUBLICATIONS   32 CITATIONS

SEE PROFILE

Eva Pilot
Maastricht University

15 PUBLICATIONS   24 CITATIONS

SEE PROFILE

Oliver Gruebner
Humboldt-Universität zu Berlin

50 PUBLICATIONS   314 CITATIONS

SEE PROFILE

Thomas Krafft
Maastricht University

104 PUBLICATIONS   509 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Air Pollution China View project

Project   Best EMS Sstem View project

Health & Place

Estimating the spatial distribution of acute undifferentiated fever (AUF) and associated risk factors using emergency call data in India. A symptom-based approach for public health surveillance

Boris Kauhl [a,*], Eva Pilot [b], Ramana Rao [c], Oliver Gruebner [d], Jürgen Schweikart [e], Thomas Krafft [a,f]

[a] Department of International Health, CAPHRI School of Public Health and Primary Care, Faculty of Health, Medicine and Life Sciences. Maastricht University, The Netherlands
[b] Department of Health, Ethics & Society, CAPHRI School of Public Health and Primary Care, Maastricht University, The Netherlands
[c] GVK Emergency Management Reseach Institute, Hyderabad, Andhra Pradesh, India
[d] Beuth University of Applied Sciences, Department III, Civil Engineering and Geoinformatics, Berlin, Germany
[e] Department of Epidemiology, Columbia University, NY, United States
[f] Institute of Environment Education and Research, Bharati Vidyapeeth University, Pune, India

ARTICLE INFO

ABSTRACT

The System for Early-warning based on Emergency Data (SEED) is a pilot project to evaluate the use of emergency call data with the main complaint acute undifferentiated fever (AUF) for syndromic surveillance in India. While spatio-temporal methods provide signals to detect potential disease outbreaks, additional information about socio-ecological exposure factors and the main population at risk is necessary for evidence-based public health interventions and future preparedness strategies. The goal of this study is to investigate whether a spatial epidemiological analysis at the ecological level provides information on urban–rural inequalities, socio-ecological exposure factors and the main population at risk for AUF. Our results displayed higher risks in rural areas with strong local variation. Household industries and proximity to forests were the main socio-ecological exposure factors and scheduled tribes were the main population at risk for AUF. These results provide additional information for syndromic surveillance and could be used for evidence-based public health interventions and future preparedness strategies.

© 2014 Published by Elsevier Ltd.

1. Introduction

The burden of disease in India is currently changing from being dominated by communicable diseases to chronic life-style related diseases. The overall burden of disease accounts for approx. 269 million disability adjusted life years (DALY) in India. Despite the epidemiological transition, communicable diseases still account for 50% of DALYs followed by 33% for non-communicable diseases and 17% for injuries (Gupte et al., 2001). Infectious and parasitic diseases are the major contributor to communicable diseases followed by respiratory infections, diarrhoeal diseases and childhood diseases (Gupte et al., 2001). Acute undifferentiated fever (AUF) is a first indicator for infectious diseases and is a major public health problem in India. The aetiology of AUF is fairly diverse and includes a wide range of infectious diseases such as dengue (Reller et al., 2012), malaria (Joshi et al., 2008), typhoid (Gasem et al., 2009), tuberculosis (Abrahamsen et al., 2013), hantavirus (Chandy et al., 2009) and Japanese encephalitis (Robertson et al., 2013).

Socio-economic disparities are a key driver not only of high rates of infectious diseases (Gupta et al., 2011, Pascual Martinez et al., 2012), especially in rural areas (Patil et al., 2002), but also of a wide range of other health problems including neonatal mortality (Kumar et al., 2013), inequalities in immunisation coverage (Lauridsen and Pradhan, 2011), mental disorders (Shidhaye and Patel, 2010) and low birth-weight (Bharati et al., 2011). The vulnerability to infectious diseases among various disadvantaged population sub-groups such as scheduled castes and scheduled tribes varies widely among and within the states of India (Raju et al., 1999), depending on the local interplay between agent, host and environmental factors (Gupte et al., 2001). A spatial epidemiological approach using Geographic Information Systems (GIS) is therefore essential to estimate the impact of socio-economic and

environmental (socio-ecological) characteristics on the incidence of infectious diseases. Such an approach has shown to deliver substantial background information for evidence-based public health interventions (Weisent et al., 2012, Haque et al., 2012, Khormi and Kumar, 2011). However, reliable and complete surveillance data is scarce in India (Aparajita and Ramanakumar, 2004, John et al., 2011), making the application of spatial epidemiological methods more challenging.

The federal structure of the Indian public health system with its variety of stakeholders and institutions, the increase of the private medical sector, the missing collaboration between the institutions and the multiplicity of vertically organised surveillance programs with their different systems of data collection complicate a uniform surveillance system (John et al., 2011). The Integrated Disease Surveillance Project (IDSP) was initiated in 2004 by the Ministry of Health and Family Welfare (MOHFW) with financial help of the World Bank and technical assistance of the World Health Organization (WHO) and the US Centers for Disease Control and Prevention (CDC). The goal of this project was to connect all district hospitals and medical colleges to establish a decentralised, state-based disease and syndromic surveillance system (Kant and Krishnan, 2010). However, this approach is not spatially inclusive as the IDSP still faces problems to include data from the private medical sector and therefore underestimates the burden of disease. The current approach to estimate the burden of disease relies on fragmentary databases derived usually from public medical facilities that serve only a small fraction of the population (John et al., 2011). The importance of including the private medical sector into disease surveillance can best be described by the following numbers: After the turn of the millennium, 67% of all hospitals, 63% of all pharmacies and 78% of all doctors were employed within the private medical sector (Patil et al., 2002). In addition, the IDSP still remains suboptimal for the control of infectious diseases. The surveillance data is often delayed, unreliable, inconsistent and the reporting rates display strong regional differences (Gaikwad et al., 2010).

The System for Early-warning based on Emergency Data (SEED) is a pilot project set up by GVK Emergency Management Research Institute (GVK EMRI), India's largest private emergency medical service provider, and GEOMED research to evaluate the use of emergency call data with the main complaint fever for syndromic surveillance of infectious diseases in India (Pilot et al., 2011; Jena et al., 2010). The project is closely linked to the European Emergency data-based System for Information on, Detection and Analysis of Risks and Threats to Health (SIDARTHa), (Office, 2014).

GVK EMRI currently operates in 14 states and 2 union territories of India, providing a chance to set up a large-scale syndromic surveillance system covering a large part of the population. The emergency call data are automatically captured using Computer Telephone Integration technology. These data are standardized, available in near real-time, spatially inclusive at fine geographic scales for the covered areas and allow the use of symptom-based data on AUF to estimate the burden of infectious diseases in areas where reliable surveillance data are not available (Joshi et al., 2008, Robertson et al., 2013).

While the general use of syndromic surveillance lies in the observation of spatial variations of common illnesses over time (Cooper et al., 2008, Horst and Coco, 2010) and the detection of potential disease outbreaks (Pilot et al., 2011; Van Den Wijngaard et al., 2010), a purely spatial, cross-sectional epidemiological analysis at the ecological level may provide additional information about socio-economic and environmental risk factors (Robertson et al., 2013, Weisent et al., 2012, Hu et al., 2012).

Infectious diseases presenting with symptoms of fever such as malaria, dengue and typhoid are driven by socio-economic, demographic and environmental characteristics (Winskill et al., 2011, Corner et al., 2013, Khormi and Kumar, 2011) and typically display higher rates in rural areas of India (Patil et al., 2002). Location-based knowledge on socio-ecological exposures and the population at risk is critical to allocate scarce financial resources (Patil et al., 2002). Such knowledge informs future preparedness strategies, for example through targeted distribution of insecticide treated bed nets.

The goal of this study is therefore to examine whether a spatial epidemiological analysis at the ecological level provides background information on the main socio-ecological exposure factors and the population at risk for evidence-based public health interventions and future preparedness strategies. Specifically, we hypothesise (i) that AUF displays higher rates in rural areas as compared to urban areas (ii) that AUF is distributed unequally across space and (iii) that AUF is associated with lower socio-economic status.

2. Methods

2.1. Study area

SEED was set up as a pilot project in three districts of Andhra Pradesh (AP), India. These three districts were selected by GVK EMRI based on their proportion of infant mortality rates, female literacy, urbanisation, proportion of reported fever and infection cases and proportion of scheduled caste and scheduled tribe population to ensure a representative sample within Andhra Pradesh (Jena et al., 2010). Srikakulam district was chosen for this study because it has the largest proportion of fever among the three selected districts. A community level household survey estimated the prevalence of fever to be 16.7%. A more detailed analysis revealed that 18% of these fever cases were attributable to malaria, 8% to typhoid and 4% to dengue and the remaining 72% to AUF (Jena et al., 2010). The district is characterised by a long coastline in the east and forested areas in the northern and north-western parts. Srikakulam has a population of 2.54 mio inhabitants according to the Census of India 2001 (Commissioner, Census Data, 2001). The smallest administrative units in rural areas of India are villages, which can be defined as areas with (i) a maximum population of 5000 inhabitants, (ii) a maximum of 75% of the male population employed in the non-agricultural sector and (iii) a maximum population density of 400 inhabitants per km^2 (Commissioner, Census Data, 2001). Mandals are the smallest administrative unit in AP for which a wide variety of population statistics are available and comprise between 27,141 and 187,132 inhabitants in Srikakulam district (Commissioner, Census Data, 2001) (Fig. 1). The district is predominantly rural and contains 11% of urban population, which is far lower than the average of 27.3% in Andhra Pradesh (Commissioner, Census Data, 2001). The literacy rate may be considered as low with only 54% as compared to 60% for the AP average. Srikakulam has a lower proportion of scheduled caste population with 9.5% as compared to the AP average of 16.0%. The proportion of scheduled tribes is slightly higher with 7.1% than the AP average of 7%.

3. Data

3.1. Outcome variable

Emergency call data with the main complaint AUF were used as indicator for infectious diseases. The emergency call data were provided by GVK EMRI and were available for the time period January 1st to December 31st, 2008. 8062 AUF calls were recorded for the year 2008 in Srikakulam district. The emergency call data were available on village level and were aggregated to mandal
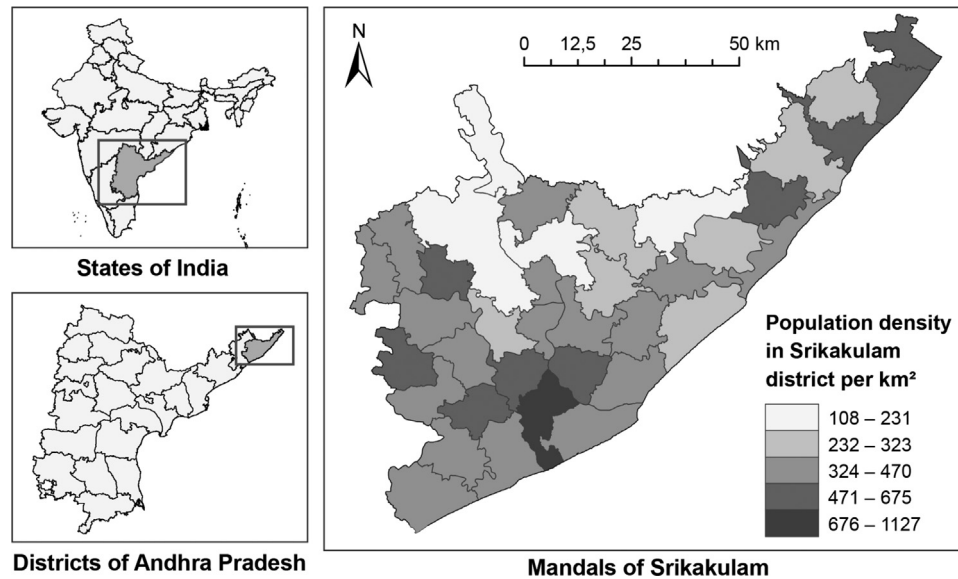
Fig. 1. Study Area, 2008.

level to be able to use population data, which were obtained from the Census of India 2001 (Commissioner, Census Data, 2001). The calculated risk expressed as the number of AUF cases for 2008 per 100,000 inhabitants for each mandal was used as the dependent variable in the regression model.

3.2. Explanatory variables

We included environmental factors associated with vector-borne diseases resulting in AUF such as rainfall (Reid et al., 2012, Alzahrani et al., 2013) and proximity to forests (Haque et al., 2011, De Castro et al., 2006). Annual Rainfall data were obtained for the year 2008 based on mandal level from the Directorate of Economics and Statistics, Hyderabad, Andhra Pradesh, India. Data on forest cover were downloaded from Open Street Map (Geofabrik, 2014). We visually checked the accuracy of the Open Street Map layer. Although not 100% accurate, we found this dataset superior than available raster datasets and sufficient for our analysis. The distance to forests was calculated as distance of the AUF emergency calls on village level to forest, averaged per mandal. To determine whether AUF follows a distinctive socio-economic gradient, we included several socio-economic variables from the Census of India 2001. An overview over all candidate explanatory variables is shown in Table 1. These variables include the sex ratio for the total population as well as for the child population (aged 0–6) measured as number of female persons per 1000 male persons. The proportion of scheduled caste and scheduled tribes represents the lowest socio-economic status since these two population groups are historically disadvantaged and have the lowest socio-economic status within the Indian society (Mohindra et al., 2006). Literacy rate contains all persons aged seven and older, who are able to read and write in any language. Literacy rate is an important predictor for understanding health-education messages and awareness of health-programs (Kumar and Quinn, 2012). Employment status was split in several categories: General work participation and proportion of main workers were included as indicator for the ability to pay for health-related costs. The variable non-workers includes persons with no personal income and is therefore an indicator for the proportion of persons unable to pay out of pocket for medical expenses. Cultivation and agricultural labour were included as potential predictors for exposure to zoonotic diseases resulting in AUF (Khan et al., 2011,

Table 1
Candidate explanatory variables. $N= =38$.

Variable	Source	Mean	SD
Children aged 0–6	Census of India	13.4%	0.9%
Sex ratio total population	Census of India	1014	38
Sex ratio child population	Census of India	969	22
Scheduled caste	Census of India	9.4%	4.1%
Scheduled tribe	Census of India	7.1%	14.6%
Literacy rate	Census of India	54%	6.7%
Work participation	Census of India	48.5%	4.3%
Main workers	Census of India	35.1%	4.9%
Cultivation labour	Census of India	23%	6.4%
Agricultural labour	Census of India	47.7%	8.8%
Household industries	Census of India	4%	1.5%
Other workers	Census of India	25.3%	12.3%
Population density	Census of India	420.6	172.9
Annual Rainfall/mandal	Dir. of Econ. and Stat.	1027 mm	331 mm
Distance to forests	Open street map	9.5 km	8.1 km

Chandy et al., 2009). Household industries are traditionally home-based and are characterised by their high level of exploitation and are another indicator of low socio-economic status (Raju et al., 1999). Other workers were included as indicator for a higher socio-economic status since this category encompasses work, which requires higher levels of education and therefore generates higher wages such as teachers, municipal servants and government employees. The variable population density was calculated as number of inhabitants per km^2. All socio-economic variables and their definitions were obtained from the Census of India 2001 (Commissioner, Census Data, 2001). Although the census data used in this study are from 2001, new data from the Census of India were not yet available on mandal level during the time of the analysis.

3.3. Analytical methods

The methodology applied in this study follows closely the recommendations of the CDC to investigate suspected clusters of cancers (National Center for Environmental Health, 2013) and has also been widely applied in a comparable manner to investigate clusters of infectious diseases (Robertson et al., 2013, Weisent et al., 2012): We created a thematic map displaying the relative

risk for each administrative unit; determined spatial clusters where the number of observed cases is higher than the expected cases; and applied a kernel density estimation to visualise the number of cases on village level. We then determined significant explanatory variables through OLS regression.

3.4. Exploratory disease mapping

To facilitate visual interpretation of the underlying disease process, the relative risk (RR) was calculated for each administrative unit. A map of the RR displays the ratio of observed to expected cases for each administrative unit and represents how much more common an event in this location is as compared to the global average (Berke, 2005). Spatial Empirical Bayes smoothing of the relative risk was considered useful in this study since the population at risk displayed a strong variation between the administrative areas. This leads to a large variance of the relative risk in areas where the underlying population is small and a small variance in areas where the underlying population is large (Lawson et al., 2000). The RR estimates were smoothed towards a local mean by using a nearest neighbour approach. The neighbours were defined as areas sharing a common edge or boundary (Waller and Gotway, 2004). We preferred a locally weighted Empirical Bayes smoothing approach over a global approach due to the occurrence of local clusters inherent in our data. The computation was carried out using the EB local function of the spdep package available in (Bivand, 2014, Team, 2013). For visualisation, the results were then imported in ESRI ArcGIS 10.1.

3.5. Local cluster detection

To determine administrative areas where the number of observed cases is significantly higher than the expected cases, the spatial scan statistic was applied to search for local clusters of elevated RR. We used the Poisson model where, under the null hypothesis, the cases of AUF follow an inhomogeneous Poisson process (Kulldorff, 1997). We selected the number of AUF cases in 2008, population from the census of India 2001 and the centroid coordinates for each mandal as necessary input data. The spatial scan statistic imposes a circular scanning window over the study area, flexibly in size and position. In this study, we evaluated clusters with 10% of the population at risk. This was done to detect spatial clusters as precisely as possible since the default setting of 50% is more likely to produce results of no practical use (Chen et al., 2008). The spatial scan statistic compares the observed and expected number of cases inside the scanning window to the area outside the scanning window. The calculation of the maximum likelihood is based on the number of observed and expected cases inside and outside the scanning window. The scanning window with the maximum likelihood and more cases than expected is the most likely cluster. The statistical significance is based on 999 Monte-Carlo replications where the null hypothesis of complete spatial randomness is rejected in this study if the p-value is less than 0.05 (Kulldorf et al., 1998). The application of the spatial scan statistic was performed in Kulldorf's SaTScan software version 9.2 (Kulldorff, 2013).

3.6. Kernel density estimation

The kernel density estimation was used as a complementary tool to visualise the spatial distribution of AUF emergency calls within the spatial clusters. The mandals to calculate the RR are fairly large spatial units and therefore mask important variations on village level. The kernel density estimation is an interpolation technique that creates a continuous surface derived from a point pattern that allows an easier identification of densely distributed features.

This is done by placing a symmetrical mathematical function over each point, the so-called kernel, which has its peak directly over the point with decreasing intensity towards the edge of the function. The distance from the point towards the edges is the bandwidth and determines the amount of smoothing inherent in the kernel density estimation (Smith and Bruce, 2008). Of the 8062 fever emergency calls in 2008, 7366 (91.4%) could be successfully matched with an already existing geodatabase, which contained the coordinates of the village centroids. These village coordinates served as input point pattern for the analysis. In this study, we chose a quartic distribution as mathematical function for the kernel and evaluated bandwidths of 1, 3 and 5 km. We found that a bandwidth of 3 km yielded the best results for our analysis. The calculation of the kernel density estimation was performed using CrimeStat III software (Levine, 2006). The results were imported in ESRI ArcGIS 10.1 and were displayed together with the layers for forest cover.

3.7. Regression analysis

The next step of our analysis was to model the influence of potential explanatory variables on the incidence of AUF. We specified our explanatory variables using following criteria: The coefficients are statistically significant and have the expected sign; the explanatory variables do not display multicollinearity and the residuals are normally distributed and are not spatially autocorrelated (Anselin, 2005). In order to achieve normality of the dependent variable, the dependent variable was transformed using a natural log-transformation (Esri, 2009, Zhou et al., 2006). To find a meaningful model, we used a data-mining tool called Exploratory Regression, which is available in ESRI ArcGIS 10.1. This tool is comparable to a step-wise regression. However, this tool identifies variable combinations in an OLS regression model that meet all requirements outlined above (Esri, 2013, Haque et al., 2012). The most parsimonious model with the lowest AIC value was used for further analysis. We then applied OLS regression in OpenGeoDa 1.2.0 (Anselin et al., 2006). The calculation of Moran's I to detect spatial autocorrelation of the residuals was based on first order queen contiguity where neighbours share a common edge or corner (Anselin, 2005).

4. Results

4.1. Difference between urban and rural risks for acute undifferentiated fever

The overall incidence of AUF was 317 per 100,000 inhabitants. Higher risks could generally be observed in purely rural areas ($RR=1.20$, 95% CI: 0.64–1.75) as compared to mandals containing urban areas ($RR=0.66$, 95% CI: 0.36–0.96).

4.2. Spatial inequalities of acute undifferentiated fever

Higher risks were concentrated in the northern parts of the district in close proximity to forests (Fig. 2). The spatial scan statistic detected two clusters. The most significant cluster was located in Seethampeta mandal ($p < 0.001$, $RR=9.7$, 1621 cases), which is characterised by a high proportion of forest cover. The second cluster consisted of the three mandals Meliaputti, Nandigam and Tekkali ($p < 0.001$, $RR=1.74$, 943 cases), which are also characterised by their high proportion of forest cover. The kernel density estimation revealed that AUF cases were concentrated in close proximity to, and within a forest in Seethampeta mandal. Especially Seethampeta village stands out with 824 AUF cases. This village contains the largest number of AUF cases per village within
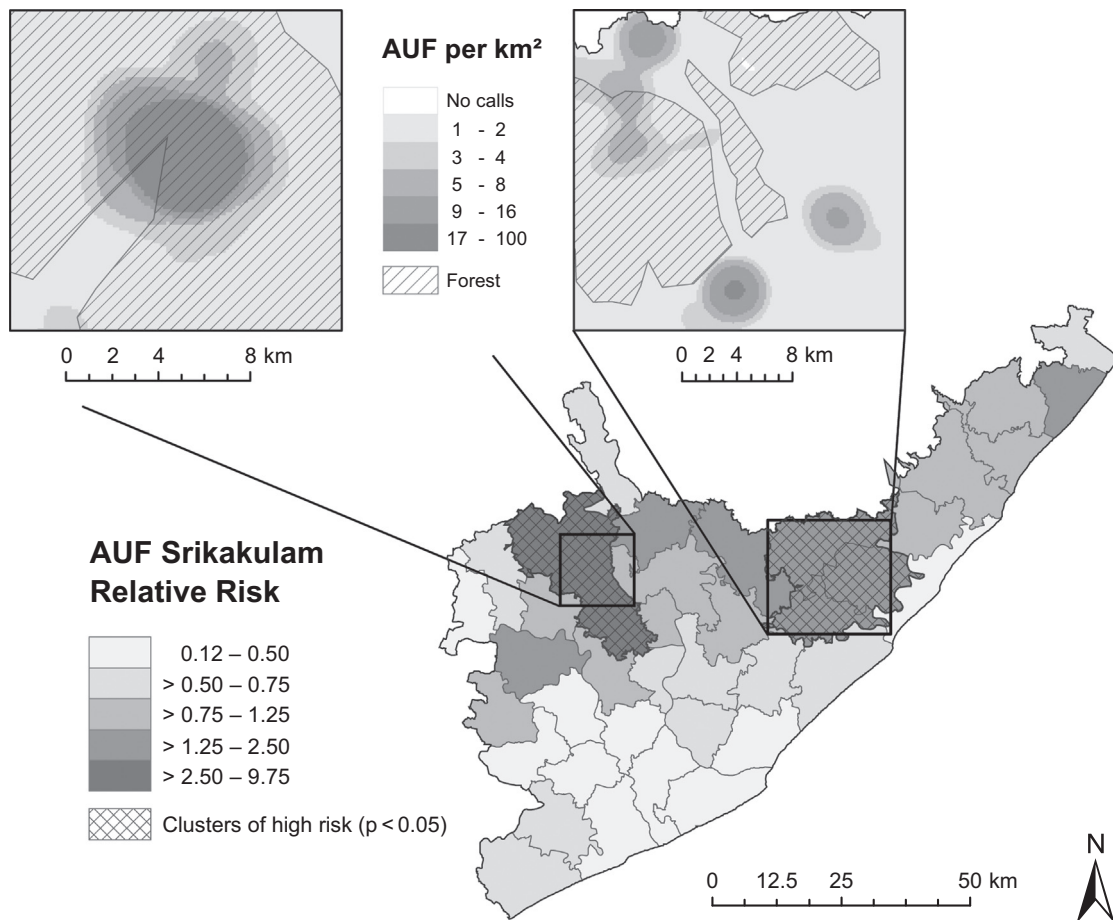
Fig. 2. Spatial clustering of AUF in Srikakulam District, 2008. Crosshatched areas indicate local clusters detected by the spatial scan-statistic.

the study area. The second largest number of AUF cases within Seethampeta mandal was observed in Pedarama village with 131 cases in close proximity to Seethampeta village. In the second spatial cluster, three concentrations stand out: The town Tekkali with 160 cases, the village Nandigam with 74 cases and the village Meliaputti with 108 cases. A spatial pattern from Meliaputti heading into the forest is visible, leading through the village Padda with 41 cases and Nelabonthu with 29 cases.

4.3. Socio-ecological exposure factors for acute undifferentiated fever

The model with the lowest AIC value and the most plausible explanation was used as the final OLS regression model, which included three explanatory variables: Percentage of scheduled tribe population, distance to forests and proportion of household industries. This model explained 66.2% of the variation in AUF emergency calls (Adj. R-squared: 0.6619). The model met all requirements for a properly specified OLS model: The model performance was overall statistically significant (F-statistic: 25.151, $p < 0.001$). The coefficients had the expected signs (Table 2) and did not display multicollinearity (multicollinearity condition number 7.408). The residuals were normally distributed (Jarque–Bera test: 1.186, $p > 0.05$) and were not spatially autocorrelated (Moran's I: 0.611, $p > 0.05$). The Lagrange multiplier tests (LM-lag and LM-error) did not show any spatial dependence (LM-lag: 0.511, $p > 0.05$; LM-error: 0.199, $p > 0.05$) implying that a spatial error model or a spatial lag model would not enhance the analysis. The coefficients revealed that the incidence of AUF is positively associated with the proportion of scheduled tribes. An increase of 1% of scheduled tribes population

Table 2
Ordinary least squares (OLS) regression coefficients.

Variable	Coefficient	Standard error	Probability
Scheduled tribes	3.04758	0.58251	< 0.001
Househ. industries	11.60569	5.50590	< 0.05
Dist. to forest (m)	− 0.04664	0.01055	< 0.001

will increase the incidence of AUF by 3%. Proportion of household industries was also positively associated with the incidence of AUF. An increase of 1% of household industries will increase the incidence of AUF by 11.6%. The distance to forest was negatively associated with the occurrence of AUF. 1 km more distance to forests will decrease the incidence of AUF by 0.047%.

By examining the spatial distribution of scheduled tribes (Fig. 3), it becomes evident that the proportion of scheduled tribes has a strong link to the incidence of AUF. Especially Seethampeta mandal stands out. In this area, the relative risk is almost 10 times as high as the district average while the proportion of scheduled tribes with 87% is almost 12 times as high as the district average. Comparable findings can also be observed for Meliaputti and Pathapatnam; the incidence of AUF and the proportion of scheduled tribes in these mandals are twice as high as compared to the district average. In contradiction, in the mandals around Srikakulam city, very low relative risks and very low proportions of scheduled tribes can be observed. The association between household industries (Fig. 4) however, are not as clear as for scheduled tribes. While it is obvious that household industries have no influence on the occurrence
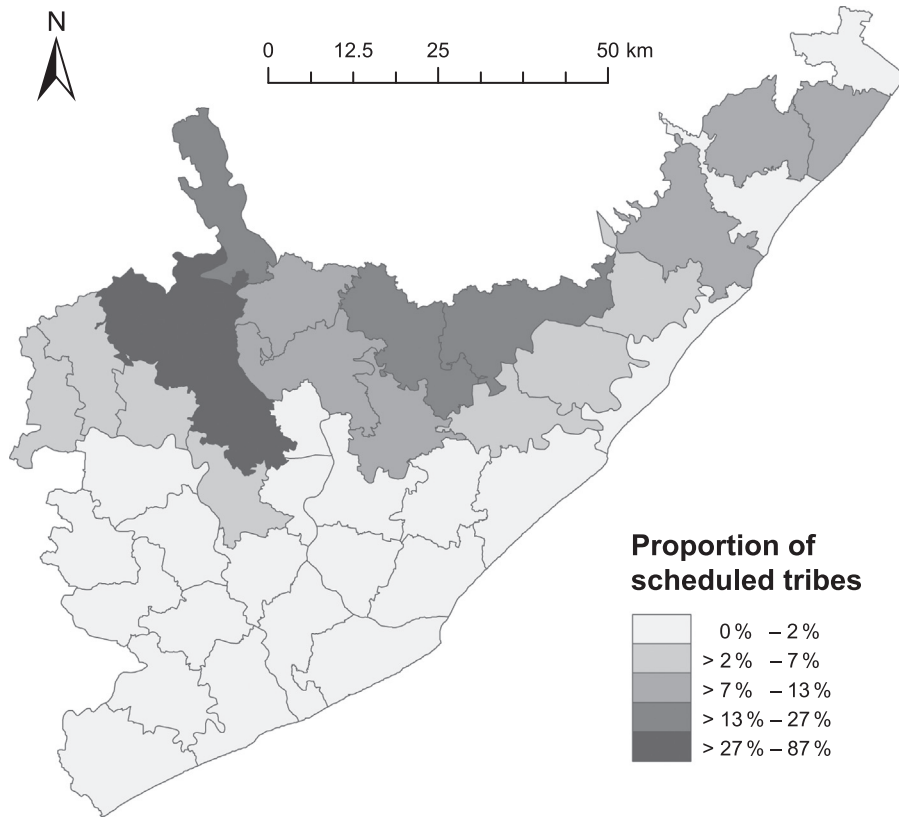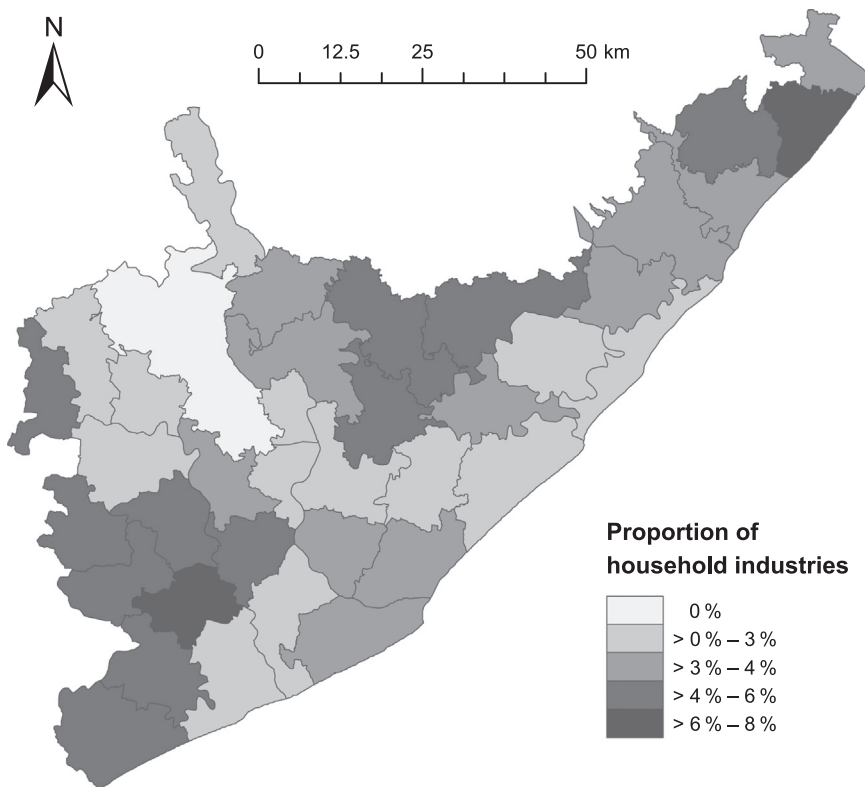
Fig. 3. Proportion of scheduled tribes.

Proportion of
scheduled tribes

0 % – 2 %
> 2 % – 7 %
> 7 % – 13 %
> 13 % – 27 %
> 27 % – 87 %

Fig. 4. Proportion of household industries

Proportion of
household industries

0 %
> 0 % – 3 %
> 3 % – 4 %
> 4 % – 6 %
> 6 % – 8 %

of AUF in Seethampeta mandal, the influence in the mandals Meliaputti, Pathapatnam and the northern mandals Kaviti and Kanchili is probably higher. The incidence of AUF shows a strong link to forested areas, especially in the most significant cluster in Seethampeta mandal but also in the second significant cluster in the mandals Meliaputti, Nandigam and Tekkali.

5. Discussion

The main findings of this study were that (i) rural areas display higher risk towards AUF as compared to urban areas (ii) that AUF is unequally distributed across mandals in Srikakulam and (iii) that scheduled tribes are the main population at risk and household industries and proximity to forests are important socio-ecological risk factors for AUF.

5.1. Higher risk of acute undifferentiated fever in rural areas

Our results suggest that the risk of AUF is higher in rural areas as compared to urban areas. These results correspond to previous findings for infectious diseases resulting in fever and corresponds well to the current health situation in rural India (Patil et al., 2002). A nationally representative survey estimated the spatial distribution of the incidence of deaths attributable to malaria in India. 90% of estimated deaths attributable to malaria occurred in rural areas and displayed strong local variation (Dhingra et al., 2010). Dengue fever in contrast, changed over time from being predominantly urban in India to gaining a strong impact in rural areas, especially in areas with dense forest (Gupta and Reddy, 2013). A wide range of other diseases such as diarrhoeal diseases and diseases carried through the air are more common in rural areas than in urban areas such as typhoid and tuberculosis and are attributable to unclean water, exposure to unhealthy living conditions and poor nutrition (Patil et al., 2002). The high correlation between the proportion of AUF emergency calls to the total emergency demand and the proportion of fever within the population as indicated through the community level household survey (Jena et al., 2010) indicates that AUF emergency calls might be a realistic indicator to estimate the burden of infectious diseases within the population. The spatial inclusiveness of these data is not only likely to show a higher incidence of infectious diseases than surveillance data would suggest but provides additionally a more reliable foundation to analyse risk factors associated with AUF than the current data of the IDSP, which is suffering from unreliable and strong regional differences in reporting rates (Gaikwad et al., 2010).

5.2. Spatial inequalities of acute undifferentiated fever

The disease mapping approach and the spatial scan statistic revealed that AUF displays strong local variation, both on mandal level as well as on village level. This variation at local level as well as the occurrence of local clusters corresponds well to previous findings analysing infectious diseases using the spatial scan statistic (Haque et al., 2009, Toan Do et al., 2013, Liebman et al., 2012). However, the full potential of the spatial scan statistic was not employed here as we followed a purely spatial approach and did not search for spatio-temporal clusters. A prospective spatio-temporal cluster detection might provide an additional value for an early-warning system based on EMS data to detect potential disease outbreaks as early as possible (Van Den Wijngaard et al., 2010).

5.3. Socio-ecological exposure factors for acute undifferentiated fever

Based on the OLS regression, we found that proportion of scheduled tribes, proportion of household industries and proximity to forests were predictors of AUF and explained 66.2% of the spatial variation of risk towards AUF. AUF risk was strongly associated with the proportion of scheduled tribes and is therefore highly correlated to the most disadvantaged population group. These results correspond to other findings from the literature showing that low socio-economic status is an important predictor of elevated rates of infectious diseases in India (Gupta et al., 2011,

Kumar et al., 2014, Sur et al., 2006), especially in rural areas (Patil et al., 2002, Pascual Martinez et al., 2012).

Indigenous population groups belong to the poorest and most disadvantaged population groups in India and research on the health of this population group is often restricted to a sample of a specific indigenous population group (Subramanian et al., 2006). Indigenous people are living often close to forest areas and are disease prone as access to health services often is limited (Balgir, 2006). Our results deliver statistical evidence for this relationship. Resulting interventions could be aimed directly at remote tribal populations to identify the underlying reasons that lead to a high vulnerability to infectious diseases. These reasons might consist of adverse distribution and poor treatment capacities of public primary healthcare facilities (Duggal, 2005), lower willingness to attend public or private health care facilities due to high out of pocket costs and loss of productivity due to absence from work (Nayar, 2007). The detection of spatial clusters might indicate areas for collecting blood samples to identify the underlying pathogens causing AUF (Phuong et al., 2006). The significant association of AUF to scheduled tribes and household industries in turn might lead to initiatives such as the provision of insecticide treated bed nets (Lengeler, 2004) or indoor residual spraying (Bousema et al., 2013).

5.4. Limitations

Our study has several limitations: The emergency calls with the main complaint AUF comprise a very broad category of potential underlying infectious diseases. This allows only an estimation of the impact of socio-economic and environmental determinants on the general incidence of certain infectious diseases but does not necessarily allow a first clue about the underlying disease itself. As shown in other studies, the complaint fever could be divided into more specific syndromes such as acute encephalitis syndrome (AES) (Joshi et al., 2008). Such an approach has shown to allow a detailed spatial analysis of landscape risk-factors associated with Japanese encephalitis (Robertson et al., 2013) and could result in more detailed knowledge about contributing ecological factors. The use of emergency calls for this study limits the explanatory power for urban areas. Due to the higher availability of transportation as well as higher availability of medical infrastructure in urban areas, the use of emergency medical services may not be the first option to use. (Saddichha et al., 2009). This highlights the need to incorporate other data-sources as well. Urban Malaria is a major public health problem in India (Kumar et al., 2014) and a strong contributor to the overall number of AUF cases in Srikakulam (Jena et al., 2010). Another limitation could be the knowledge of GVK EMRI's 108 toll free emergency hotline. We were unable to verify if the service is equally popular within the district or if there are any notable spatial gaps of advertisement. Potentially, this could lead in areas with high advertisement to more frequent use and in areas with low advertisement to an under-utilization of this service. It would be interesting to compare the results of our analysis with results based on laboratory confirmed cases of infectious diseases to see whether our results differ widely from results conducted using laboratory confirmed cases. However, given the current scenario of disease surveillance in India, such a comparison is not possible (John et al., 2011). The number of explanatory variables available from the Census of India on mandal level was very limited. We would have favoured to include different age groups as additional explanatory variables to analyse which age group is most at risk. In addition, other important variables such as bed-net use, housing materials, accessibility to health care providers and distance to water bodies as indicator for a potential vulnerability to vector-borne diseases (Haque et al., 2012) were not available for this study. The administrative units

we used in this study are fairly large areas. Although we displayed the number of cases on village level using a kernel density estimation, we could not display the incidence rate or create a spatial regression model on this scale since the necessary census data were not available on village level during the time of the analysis. This limitation decreases the explanatory power of the kernel density estimation. In addition, we would have favoured a Geographically Weighted Regression (GWR) to account for spatial heterogeneity. However, since our study area consisted only of 38 administrative units, the results would have been unreliable. Páez et al. point out that the use of GWR for small datasets with less than 160 administrative is not advisable (Páez et al., 2011). Current studies benefitting from the application of GWR usually focus on fairly large datasets (Tsai, 2013, Weisent et al., 2012, Haque et al., 2012, Hu et al., 2012). This limitation underlines, that future research on risk factors should focus on analysing AUF emergency calls on larger areas such as whole states to be able to capture spatial heterogeneity of socio-economic and environmental determinants within regression models. Such an approach could enhance the use of symptom-based data to explain the range of contributing factors to AUF.

6. Conclusions

We used EMS data with the main complaint acute undifferentiated fever as indicator for infectious diseases and linked AUF to socio-ecological exposure factors.

Our results display that the spatial distribution of AUF follows closely the current scenario of infectious diseases in India as it reflects a higher vulnerability to fever in rural areas, spatial heterogeneity at local levels and a strong association with lower socio-economic status. This in turn highlights the value of AUF emergency calls to monitor the spatial distribution of infectious diseases in areas where reliable surveillance data are not available. In addition, our approach shows that an epidemiological analysis at the ecological level using emergency call data could be used to identify main socio-ecological exposure factors and the main population at risk. These results might be relevant for future preparedness strategies and targeted, evidence-based public health interventions and provide additional information for syndromic surveillance. Our approach also stresses the importance and possibilities of including private medical institutions in surveillance activities. We hypothesise that our approach is useful not only for Srikakulam district, but also could be an effective way of guiding evidence-based public health interventions and future preparedness strategies in India where spatial EMS data are available.

Acknowledgement

References

Abrahamsen, S.K., Haugen, C.N., Rupali, P., Mathai, D., Langeland, N., Eide, G.E., Morch, K., 2013. Fever in the tropics: aetiology and case-fatality - a prospective observational study in a tertiary care hospital in South India. BMC Infect. Dis. 13, 355.

Alzahrani, A.G., Al Mazroa, M.A., Alrabeah, A.M., Ibrahim, A.M., Mokdad, A.H., Memish, Z.A., 2013. Geographical distribution and spatio-temporal patterns of

dengue cases in jeddah governorate from 2006–2008. Trans. R. Soc. Trop. Med. Hyg. 107, 23–29.

Anselin, L. 2005. Exploring Spatial Data with GeoDaTM: A Workbook [Online]. Urbana, Illinois, USA: Spatial Analysis Laboratory, Department of Geography, University of Illinois at Urbana-Champaign. Available: ⟨https://geodacenter.asu.edu/system/files/geodaworkbook.pdf⟩ (accessed 2. 03. 14.).

Anselin, L., Syabri, I., Kho, Y., 2006. GeoDa: an introduction to spatial data analysis. Geogr. Anal. 38, 5–22.

Aparajita, C., Ramanakumar, A., 2004. Burden of disease in rural India: an analysis through cause of death. Internet J. Third World Med. 2 (2).

Balgir, R. Tribal health problems, disease burden and ameliorative challenges in tribal communities with special emphasis on tribes of Orissa. In: Proceedings of National Symposium on "Tribal Health", 19th–20th October, 2006, pp. 161–176.

Berke, O., 2005. Exploratory spatial relative risk mapping. Prev. Vet. Med. 71, 173–182.

Bharati, P., Pal, M., Bandyopadhyay, M., Bhakta, A., Chakraborty, S., Bharati, P., 2011. Prevalence and causes of low birth weight in India. Malays. J. Nutr. 17, 301–313.

Bivand, R. 2014. spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-71. [Online]. (accessed 12. 02. 14.).

Bousema, T., Stevenson, J., Baidjoe, A., Stresman, G., Griffin, J.T., Kleinschmidt, I., Remarque, E.J., Vulule, J., Bayoh, N., Laserson, K., Desai, M., Sauerwein, R., Drakeley, C., Cox, J., 2013. The impact of hotspot-targeted interventions on malaria transmission: study protocol for a cluster-randomized controlled trial. Trials 14, 36.

Chandy, S., Yoshimatsu, K., Boorugu, H.K., Chrispal, A., Thomas, K., Peedicayil, A., Abraham, P., Arikawa, J., Sridharan, G., 2009. Acute febrile illness caused by hantavirus: serological and molecular evidence from India. Trans. R. Soc. Trop. Med. Hyg. 103, 407–412.

Chen, J., Roth, R.E., Naito, A.T., Lengerich, E.J., Maceachren, A.M., 2008. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. Int. J. Health Geogr. 7, 57.

Commissioner, R.G.C. 2001. Census Data 2001 / Metadata [Online]. New Delhi, India: Ministry of Home Affairs, Government of India. Available: http://censusindia.gov.in/ (accessed 13. 03. 14.).

Cooper, D.L., Smith, G.E., Regan, M., Large, S., Groenewegen, P.P., 2008. Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data. BMC Med. 6 (16).

Corner, R.J., Dewan, A.M., Hashizume, M., 2013. Modelling typhoid risk in Dhaka metropolitan area of Bangladesh: the role of socio-economic and environmental factors. Int. J. Health Geogr. 12, 13.

De Castro, M.C., Monte-Mor, R.L., Sawyer, D.O., Singer, B.H., 2006. Malaria risk on the amazon frontier. Proc. Natl. Acad. Sci. USA 103, 2452–2457.

Dhingra, N., Jha, P., Sharma, V.P., Cohen, A.A., Jotkar, R.M., Rodriguez, P.S., Bassani, D.G., Suraweera, W., Laxminarayan, R., Peto, R., Million Death Study Collaborators, 2010. Adult and child malaria mortality in India: a nationally representative mortality survey. Lancet 376, 1768–1774.

Duggal, R., 2005. Public Health Expenditures, Investment and Financing Under the Shadow of a Growing Private Sector. Review of Healthcare in India, CEHAT, Mumbai.

Esri. 2009. Regression analysis basics [Online]. Available: http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Regression_analysis_basics (accessed 11. 03. 14.).

Esri. 2013. How Exploratory Regression works [Online]. Available: http://resources.arcgis.com/en/help/main/10.1/index.html -//005p00000054000000 (accessed 06. 02. 14.).

Gaikwad, A., Oruganti, R., Singh, V., Anchala, R., Rao, B.M., Ramswaropp, 2010. Reporting pattern in integrated disease surveillance project (IDSP) in Andhra Pradesh. Indian Emerg. J. 5, 13–16.

Gasem, M.H., Wagenaar, J.F., Goris, M.G., Adi, M.S., Isbandrio, B.B., Hartskeerl, R.A., Rolain, J.M., Raoult, D., Van Gorp, E.C., 2009. Murine typhus and leptospirosis as causes of acute undifferentiated fever, Indonesia. Emerg. Infect. Dis. 15, 975–977.

Geofabrik. 2014. India [Online]. Open Street Map. Available: http://download.geofabrik.de/asia/india.html (accessed 22. 02. 14.).

Gupta, B., Reddy, B.P., 2013. Fight against dengue in India: progresses and challenges. Parasitol. Res. 112, 1367–1378.

Gupta, S., Shenoy, V.P., Mukhopadhyay, C., Bairy, I., Muralidharan, S., 2011. Role of risk factors and socio-economic status in pulmonary tuberculosis: a search for the root cause in patients in a tertiary care hospital, South India. Trop. Med. Int. Health 16, 74–78.

Gupte, M.D., Ramachandran, V., Mutatkar, R.K., 2001. Epidemiological profile of India: historical and contemporary perspectives. J. Biosci. 26, 437–464.

Haque, U., Huda, M., Hossain, A., Ahmed, S.M., Moniruzzaman, M., Haque, R., 2009. Spatial malaria epidemiology in Bangladeshi highlands. Malar. J. 8, 185.

Haque, U., Scott, L.M., Hashizume, M., Fisher, E., Haque, R., Yamamoto, T., Glass, G.E., 2012. Modelling malaria treatment practices in Bangladesh using spatial statistics. Malar. J. 11, 63.

Haque, U., Sunahara, T., Hashizume, M., Shields, T., Yamamoto, T., Haque, R., Glass, G.E., 2011. Malaria prevalence, risk factors and spatial distribution in a hilly forest area of Bangladesh. PLoS One 6, e18908.

Horst, M.A., Coco, A.S., 2010. Observing the spread of common illnesses through a community: using Geographic Information Systems (GIS) for surveillance. J. Am. Board Fam. Med. 23, 32–41.

Hu, M., Li, Z., Wang, J., Jia, L., Liao, Y., Lai, S., Guo, Y., Zhao, D., Yang, W., 2012. Determinants of the incidence of hand, foot and mouth disease in China using geographically weighted regression models. PLoS One 7, e38978.

Jena, B., Prasad, M., Murthy, S., Ramanarao, G., 2010. Demand pattern for medical emergency services for infectious diseases in Andhra Pradesh-a geo-spatial temporal analysis of fever cases. Indian Emerg. J. 5, 5–8.

John, T.J., Dandona, L., Sharma, V.P., Kakkar, M., 2011. Continuing challenge of infectious diseases in India. Lancet 377, 252–269.

Joshi, R., Colford Jr., J.M., Reingold, A.L., Kalantri, S., 2008. Nonmalarial acute undifferentiated fever in a rural hospital in central India: diagnostic uncertainty and overtreatment with antimalarial agents. Am. J. Trop. Med. Hyg. 78, 393–399.

Kant, L., Krishnan, S.K., 2010. Information and communication technology in disease surveillance, India: a case study. BMC Public Health 10, S11.

Khan, S.M., Debnath, C., Pramanik, A.K., Xiao, L., Nozaki, T., Ganguly, S., 2011. Molecular evidence for zoonotic transmission of *Giardia duodenalis* among dairy farm workers in West Bengal, India. Vet. Parasitol. 178, 342–345.

Khormi, H.M., Kumar, L., 2011. Modeling dengue fever risk based on socioeconomic parameters, nationality and age groups: GIS and remote sensing based case study. Sci. Total Environ. 409, 4713–4719.

Kulldorf, M., Athas, W.F., Feuer, E.J., Miller, B.A., Key, C.R., 1998. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. Am. J. Public Health 88, 1377–1380.

Kulldorff, M., 1997. A spatial scan statistic. Commun. Stat.-Theory Methods 26, 1481–1496.

Kulldorff, M. 2013. SaTScanTM User Guide for version 9.2 [Online]. (accessed 08. 09. 13.).

Kumar, C., Singh, P.K., Rai, R.K., Singh, L., 2013. Early neonatal mortality in India, 1990–2006. J. Community Health 38, 120–130.

Kumar, D.S., Andimuthu, R., Rajan, R., Venkatesan, M.S., 2014. Spatial trend, environmental and socioeconomic factors associated with malaria prevalence in Chennai. Malar. J. 13, 14.

Kumar, S., Quinn, S.C., 2012. Existing health inequalities in India: informing preparedness planning for an influenza pandemic. Health Policy Plan. 27, 516–526.

Lauridsen, J., Pradhan, J., 2011. Socio-economic inequality of immunization coverage in India. Health Econ. Rev. 1, 11.

Lawson, A.B., Biggeri, A.B., Boehning, D., Lesaffre, E., Viel, J.F., Clark, A., Schlattmann, P., Divino, F., 2000. Disease mapping models: an empirical evaluation. Disease mapping collaborative group. Stat. Med. 19, 2217–2241.

Lengeler, C., 2004. Insecticide-treated bed nets and curtains for preventing malaria. Cochrane Database Systematic Reviews 2, http://dx.doi.org/10.1002/14651858. CD000363, Art no.:CD000363.

Levine, N., 2006. Crime mapping and the crimestat program. Geogr. Anal. 38, 41–56.

Liebman, K.A., Stoddard, S.T., Morrison, A.C., Rocha, C., Minnick, S., Sihuincha, M., Russell, K.L., Olson, J.G., Blair, P.J., Watts, D.M., Kochel, T., Scott, T.W., 2012. Spatial dimensions of dengue virus transmission across interepidemic and epidemic periods in Iquitos, Peru (1999–2003). PLoS Negl. Trop. Dis. 6, e1472.

Mohindra, K.S., Haddad, S., Narayana, D., 2006. Women's health in a rural community in Kerala, India: do caste and socioeconomic position matter? J. Epidemiol. Community Health 60, 1020–1026.

(MMWR Recommendations and Reports, 62). , pp. 1–24.

Nayar, K., 2007. Social exclusion, caste & health: a review based on the social determinants framework. Indian J. Med. Res. 126, 355.

Office E.-N. C. , 2014. European Emergency Data-based Syndromic Surveillance System [Online]. European Emergency Data Group - Research Network Coordination Office. Available: http://sidartha.eu/ (accessed 11. 03. 14.).

Páez, A., Farber, S., Wheeler, D., 2011. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. Environ. Plan.-Part A 43, 2992.

Pascual Martinez, F., Picado, A., Roddy, P., Palma, P., 2012. Low castes have poor access to visceral leishmaniasis treatment in Bihar, India. Trop. Med. Int. Health 17, 666–673.

Patil, A.V., Somasundaram, K.V., Goyal, R.C., 2002. Current health scenario in rural India. Aust. J. Rural Health 10, 129–135.

Phuong, H.L., De Vries, P.J., Nga, T.T., Giao, P.T., Hung, L.Q., Binh, T.Q., Nam, N.V., Nagelkerke, N., Kager, P.A., 2006. Dengue as a cause of acute undifferentiated fever in Vietnam. BMC Infect. Dis. 6, 123.

Pilot, E., Jena, B., Krafft, T., Ziemann, A., Rao, R., Kauhl, B., Pardeshi, RR., Salvi, S., 2011. New approaches to disease surveillance and early outbreak detection in India. In: Tripathi, N.K., Joshi, P.K., Mehmood, H. Proceedings of the Fourth International Conference on Health GIS, NewDelhi, India, 5-6 August 2011. Managing Health Geospatially, Teri, pp. 240–246.

Raju, S., Atkins, P.J., Kumar, N., Townsend, J.G., Corbridge, S., Duvvury, N., Evans, I., Harriss, B., Kumar, S., Oughton, E.A., 1999. Atlas of Women and Men in IndiaZubaan Books, New Delhi, India.

Reid, H.L., Haque, U., Roy, S., Islam, N., Clements, A.C., 2012. Characterizing the spatial and temporal variation of malaria incidence in Bangladesh, 2007. Malar. J. 11, 170.

Reller, M.E., Bodinayake, C., Nagahawatte, A., Devasiri, V., Kodikara-Arachichi, W., Strouse, J.J., Broadwater, A., Ostbye, T., De Silva, A., Woods, C.W., 2012. Unsuspected dengue and acute febrile illness in rural and semi-urban southern Sri Lanka. Emerg. Infect. Dis. 18, 256–263.

Robertson, C., Pant, D.K., Joshi, D.D., Sharma, M., Dahal, M., Stephen, C., 2013. Comparative spatial dynamics of Japanese encephalitis and acute encephalitis syndrome in Nepal. PLoS One 8, e66168.

Saddichha, S., Saxena, M.K., Vibha, P., Methuku, M., 2009. Neurological emergencies in India–lessons learnt and strategies to improve outcomes. Neuroepidemiology 33, 280–285.

Shidhaye, R., Patel, V., 2010. Association of socio-economic, gender and health factors with common mental disorders in women: a population-based study of 5703 married rural women in India. Int. J. Epidemiol. 39, 1510–1521.

Smith, S.C. & Bruce, C.W. 2008. CrimeStat III User Workbook [Online]. Washington, DC: The National Institute of Justice. Available: ⟨https://http://www.icpsr.umich. edu/CrimeStat/workbook/CrimeStat_Workbook.pdf⟩ (accessed 12. 03. 14.).

Subramanian, S.V., Davey Smith, G., Subramanyam, M., 2006. Indigenous health and socioeconomic status in India. PLOS Med. 3, e421.

Sur, D., Von Seidlein, L., Manna, B., Dutta, S., Deb, A.K., Sarkar, B.L., Kanungo, S., Deen, J.L., Ali, M., Kim, D.R., Gupta, V.K., Ochiai, R.L., Tsuzuki, A., Acosta, C.J., Clemens, J.D., Bhattacharya, S.K., 2006. The malaria and typhoid fever burden in the slums of Kolkata, India: data from a prospective community-based study. Trans. R. Soc. Trop. Med. Hyg. 100, 725–733.

Team, R.C. 2013. A language and environment for statistical computing. [Online]. Vienna, Austria: R Foundation for Statistical Computing. Available: ⟨http://www.r-project.org⟩ (accessed 14. 01. 14.).

Toan Do, T.T., Hu, W., Quang Thai, P., Hoat, L.N., Wright, P., Martens, P., 2013. Hot spot detection and spatio-temporal dispersion of dengue fever in Hanoi, Vietnam. Glob. Health Action 6, 18632.

Tsai, P.J., 2013. Scrub typhus and comparisons of four main ethnic communities in taiwan in 2004 versus 2008 using geographically weighted regression. Glob. J. Health Sci. 5, 101–114.

Van Den Wijngaard, C.C., Van Asten, L., Van Pelt, W., Doornbos, G., Nagelkerke, N.J., Donker, G.A., Van Der Hoek, W., Koopmans, M.P., 2010. Syndromic surveillance for local outbreaks of lower-respiratory infections: would it work? PLoS One 5, e10406.

Waller, L., Gotway, C., 2004. Applied Spatial Statistics for Public Health Data. John Wiley and Sons, Inc, Hoboken, NJ.

Weisent, J., Rohrbach, B., Dunn, J.R., Odoi, A., 2012. Socioeconomic determinants of geographic disparities in campylobacteriosis risk: a comparison of global and local modeling approaches. Int. J. Health Geogr. 11, 45.

Winskill, P., Rowland, M., Mtove, G., Malima, R.C., Kirby, M.J., 2011. Malaria risk factors in north-east Tanzania. Malar. J. 10, 98.

Zhou, Y., Hallisey, E.J., Freymann, G.R., 2006. Identifying perinatal risk factors for infant maltreatment: an ecological approach. Int. J. Health Geogr. 5, 53.