G Model
CEGH 175 No. of Pages 6

ARTICLE IN PRESS

Clinical Epidemiology and Global Health xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Clinical Epidemiology and Global Health

journal homepage: www.elsevier.com/locate/cegh

Original Article

Modelling spatial patterns of misaligned disease data: An application on measles incidence in Namibia

D. Ntirampeba[a,*], I. Neema[b], L.N. Kazembe[c]

[a] *Department of Mathematics and Statistics, Namibia University of Science and Technology, Windhoek 264, Namibia*
[b] *Namibia Statistics Agency (NSA), Windhoek 264, Namibia*
[c] *Department of Statistics and Population Studies, University of Namibia, Windhoek 264, Namibia*

ARTICLE INFO

Article history:
Received 18 November 2016
Accepted 24 January 2017
Available online xxx

Keywords:
Hierarchical Bayesian model
Misalignment
Measles
Namibia
Spatial modelling

ABSTRACT

Introduction: Quite often disease data are available in aggregated formats mostly to maintain confidentiality. This leads to a misalignment problem when the goal is to analyze risk at a different level of spatial resolution different from the original administrative level where data were available.
Objective: To estimate and map the risk of measles at a sub-region level in Namibia using data obtained at a regional level.
Methods: Using measles data from Namibia for the period 2005–2014, both multi-step and direct approaches were applied to correct for misalignment. Subsequently, ecological Bayesian regression models were fitted and compared.
Results: Results show that the variables standardized birth rate, counts of measles cases for previous year, unemployment rate and proportion of vaccinated children against measles by age 12 months were significant determinants of measles risk. Constituencies having elevated measles risk were identified mostly in the northern corridor with Angola.
Conclusion: We recommend that relevant authorities should make geographical target intervention and redesign prevention and control strategies based on these findings.

1. Introduction

Measles is a disease caused by a highly contagious human pathogen that belongs to the Paramyxoviridae family.[1] The disease spreads through coughing, sneezing, near contact or direct contact with infected nasal or throat secretions. It has an incubation period located between 9 and 12 days and an infectivity period located between 4 and 9 days.[2] Deaths due to measles are quite common among malnourished children and people whose immune system has been weakened by diseases that include human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS). Measles leads to other complications such as blindness, brain swelling (encephalitis), diarrhoea, ear infections and respiratory infection such as pneumonia. High death rates are commonly registered in developing countries with low per capita income and poor health service systems.[3]

Worldwide, measles is ranked among the leading causes of mortality especially among children in developing countries. For instance, in 2013, about 145,700 deaths were recorded.[4] Until now, there is no antiviral treatment for measles virus. Thus far, measles vaccination and supportive care that includes good nutrition and adequate fluid intake have been used to fight measles.[4] However, reduction of global funding by the governments and partners has largely affected the immunization campaigns, which hampered efforts for a complete elimination of measles.[3] Consequently, measles cases are still reported in many countries, with Angola, Ethiopia, Namibia, Bosnia and Herzegovina, Georgia, Sri Lanka and Philippines ranked among the top ten countries with high annualized measles incidence per 100,000 inhabitants in 2014.[5]

In Namibia, as in many countries, diseases surveillance data are often analyzed in the form of aggregated data at health district or regional level because of confidentiality issues. However, health decisions might be needed at lower political boundaries such as constituencies. Nevertheless, direct inference at such lower level made on the basis of regionally aggregated data may lead to spatial misalignment problem.[6]

In brief, spatial misalignment appears through various processes. The first process is when the purpose of the analysis is to make inference about new points based on available information at different points or locations. This is known as point-to-point change of support. The second is when a researcher might be

* Corresponding author.
E-mail address: dntirampeba@Nust.Na (D. Ntirampeba).

http://dx.doi.org/10.1016/j.cegh.2017.01.002
2213-3984/

Please cite this article in press as: D. Ntirampeba, et al., Modelling spatial patterns of misaligned disease data: An application on measles incidence in Namibia, Clin Epidemiol Glob Health (2017), http://dx.doi.org/10.1016/j.cegh.2017.01.002

interested at predicting values at blocks level using information available at point level. This is called the point-to-block change of support. In the third process, one might seek to make inference from block values to point level, and is referred as the block to point change of support.[6] In this scenario, it is inappropriate to infer about the relationships between variables at individual level using information observed at area level, as the accuracy at area and point levels is not a one-to-one relationship. This challenge is referred to as ecological fallacy. Fourth, spatial misalignment can arise when the purpose of the spatial analysis is the interpolation at new aggregation level that is different from a level where data were observed. Scholars refer this to as modifiable area unit problem.

Various methods for resolving misalignment have been proposed.[7–13] For instance, methods have been applied to downscale the distributions of data from coarse to fine grain that include direct method, point sampling method and hierarchical Bayesian method. These methods have been adopted to deal with this scenario of misalignment.[8–10] Other techniques have been developed that deal with spatial misalignment that arises when the response variable is available at bigger irregular-shaped area units and covariates are available at smaller fine grids,[11] in which a multi-step approach has been applied. In the case where misalignment occurs with non-nested overlapping grids, hierarchical Bayesian approaches have been employed.[6,12] Recently, the later has been extensively applied as it permits to derive posterior predictive distributions for both parameters and epidemiological outcome of interest. It is also suitable when dealing with multiple sources of uncertainty and enables to incorporate additional sources of information in the form of prior knowledge.[13]

The aim of this study is two-fold. First, the study aimed to identify constituencies (sub-regions) in Namibia that have elevated risk of measles and also to visualize smoothed patterns of risk of measles. Second, the study aimed to determine factors associated with the risk of measles in Namibia.

2. Methods

2.1. Sources of data

Measles cases were abstracted from the health management information system (HMIS) database within the Ministry of Health and Social Services (MoHSS) in Namibia. The database included all suspected measles cases from which positive cases were extracted. Any patient consulting a health facility becomes a suspected case if the patient is diagnosed with fever and generalized maculopular rash lasting for three days or longer and cough, coryza or conjunctivitis. Such a case will be investigated and adequate blood specimen is collected and examined at the Namibia Institute of Pathology (NIP). If the blood specimen is found to have serological confirmation of a recent virus infection, the case is classified as laboratory confirmed. However, there are other cases wherein blood specimens are not taken for serological, but are linked to laboratory confirmed cases. Such cases are known as epidemiologically confirmed. A suspected case is discarded if it has been completely investigated or the blood specimen is declared by NIP as not having serological evidence of recent measles virus. This determination of a measles case is based on World Health Organization's (WHO) standard definition, which considers a measles case as either an epidemiologically confirmed case or a laboratory confirmed case.[4,5] Although the HMIS database has information from 2001 to partly 2015, the 2005–2014 period provided consistent information for the entire country, and hence only data from this period were considered in this study.

Covariates used in this study were obtained from the 2011 Namibia population and housing census (NPHC) and the 2013 Namibia demographic health survey (NDHS). The variables include proxies of social mixing patterns (average household size and proportions of children attending pre-primary and schools), unemployment rates and birth rates. Table 1 provides a list of all variables used in the analysis, as identified through literature.[2,14–19] Shapefiles that defined the administrative boundary maps were also obtained from the Namibia Statistics Agency (NSA). Although the administrative boundaries have changed over time, in this study we have used the 2011 administrative boundaries that match with variables derived from the 2011 NPHC.

2.2. Statistical methods

Each of the 13 regions in Namibia is sub-divided into constituencies, giving a total of 107 sub-regions. The counts of measles cases are available at 13 regions, of which our aim is to estimate the risk at constituency level. This introduces the problem of misalignment in the analysis. To overcome misalignment, two approaches (multi-step and direct methods) are used.

2.2.1. Multi-step approach

This method allows the allocation of region/district disease totals to constituency proportional to the constituency area or population. However, the areal area proportional allocation method assumes that population is uniformly distributed throughout the entire area. Namibia is a semi-desert country, and its population is not spread uniformly throughout the territory; rather, people are living in towns or settlements. It has been shown that the measles infection is proportional to the size of the population in each location.[2] Thus, the population proportional allocation was applied. The steps of multi-step method are as follows:

(i) Overlay constituencies on regions. This enables to determine exactly what proportion of a given constituency is susceptible or infected by measles.
(ii) Find all total values of measles cases for all constituencies. The computation of these values is based on population proportional allocation concept, as it has been shown to be more appropriate relative to areal proportional allocation for infectious diseases.[2] This is formulated as follows: $y_{ik} = \frac{P_{ik}}{P_k} Y_k$, where y_{ik} is the number of measles cases in the ith

Table 1
Description of variables considered for the analysis.

Variable	Variable name
1	Standardized average household size
2	Counts of measles for previous year (2004)
3	Unemployment rates
4	Standardized birth rates
5	Proportion of children attending schools
6	Proportion of vaccinated children against measles by age 12 months
7	Proportion of children attending pre-primary