Spatial and Spatio-temporal Epidemiology 9 (2014) 23–35

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

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

A space-time point process model for analyzing and predicting case patterns of diarrheal disease in northwestern Ecuador

Jaeil Ahn [a,*], Timothy D. Johnson [b], Darlene Bhavnani [c], Joseph N.S. Eisenberg [c], Bhramar Mukherjee [b]

[a] Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University, Washington, DC 20057, USA
[b] Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109 USA
[c] Department of Epidemiology, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 3 July 2013
Revised 3 December 2013
Accepted 1 February 2014
Available online 13 March 2014

Keywords:
Diarrheal disease
Kriging
LGCP
Prediction
Poisson process
Spatio-temporal

ABSTRACT

We consider modeling case-patterns under a complex spatial and longitudinal sampling design as conducted via a serial case–control study of diarrheal disease in northwestern Ecuador. We build a two-stage space-time model to understand the role of spatially and temporally referenced covariates that reflect social and natural environments in the sampled region, after accounting for unmeasured residual heterogeneities. All diarrheal case events are collected from 21 sampled communities in Esmeraldes province in Ecuador, during seven sampling cycles from 2003 to 2008. The region of interest comprises 158 communities along a river basin. Prediction of case counts at unsampled communities at a future time is of interest along with estimation of risk-related parameters. We propose a computationally feasible two-stage Bayesian approach to estimate the risk-related parameters and conduct predictive inference. We first apply the log Gaussian Cox process (LGCP), commonly used to model spatial clustering of point patterns, to accommodate temporal variation within the sampled communities. Prediction of the number of cases at unsampled communities at a future time is obtained by a disease mapping model conditional on the expected case counts from Stage I.

1. Introduction

Diarrhea is the second leading cause of death in children under five years of age (The global burden of disease: 2004 update, 2008). According to the World Health Organization ("Diarrhoeal disease", 2012), diarrheal deaths exceed the combined death toll due to AIDS, tuberculosis and malaria, largely because of the high death rate in developing countries: approximately 2.5 million deaths occur each year (Kosek et al., 2003). Previous epidemiological studies (Curtis and Cairncross, 2003;Checkley et al., 2004; Barreto et al., 2007) attest that this high prevalence is largely attributable to individual risks factors such as poor hygiene, food contamination, low socio-economic status and community-associated factors such as inferior water quality and poor or non-existent sanitation systems. Eisenberg et al. (2006) argue that future epidemiological studies should move towards a more hierarchical approach to understand the impact of changes in community level factors, household factors and individual level factors that may underlie the biological or social causes of diarrhea.

The ECODESS (Ecologia, Desarrollo, Salud, y Sociedad) study is designed to further our understanding of the underlying causal process of diarrheal transmission involving social and ecological factors such as road construction, social networks, sanitation and other confounding factors ("Environmental Change and Diarrheal Disease: A Natural

* Corresponding author. Tel.: +1 2026870197.
E-mail address: ja1030@georgetown.edu (J. Ahn).

http://dx.doi.org/10.1016/j.sste.2014.02.001
1877-5845/

Experiment", 2012)). The study design and initial findings from this study are described in Eisenberg et al., 2006). The investigators selected 21 communities in the Esmeraldes province of northwestern Ecuador for this study. The study region sits at the southern end of the Chocó rainforest and is classified as a biodiversity hotspot with high levels of endemic plant and animal species. Since the 1980s, the region has seen high rates of deforestation. The region has 158 communities located along one of three river systems (Rio Cayapas, Rio Santiago and Rio Onzole), which all drain towards Borbón, the region's population center (Fig. 1). Borbón and the surrounding region have recently undergone dramatic changes due to the construction of a new highway that has encouraged the movement of people, commercial goods and information in and out of the region.

Twenty one communities were selected by a block randomized design using location, size, population and relative distance to Borbón. Within the sampled communities, all households were enrolled in the study with a 98% participation rate. The right panel of Fig. 1 depicts the locations of all 158 communities relative to Borbón. For confidentiality reasons, we use numbers rather than the actual names of the communities throughout the paper. The ECODESS research team visited each sampled community semi-annually or annually, on a rotating basis from the beginning of August 2003 to March 2008, for a total of 7 visits, or cycles. The researchers interviewed each household every morning, for fifteen consecutive days, and identified all diarrheal cases. A case was defined as an individual having three or more loose stools in a 24-h period. Cases, as well as geographical (GIS) coordinates of the household, were recorded. Demographic data such as age, gender and sanitation were also collected. Control data were also collected but not used in this analysis.

Using a remoteness metric, based on time and cost of travel to Borbón, Eisenberg et al. (2006) found that more remote communities have a lower occurrence of the disease and this association is statistically significant. Bates et al. (2007) examined the role of social networks on disease transmission and found an association between a single summary measure of social network and disease incidence. These findings are consistent with Gushulak and MacPherson (2004) who demonstrated that remote communities, known to have lower immigration and emigration rates, have lower diarrheal transmission rates. Levy et al. (2009) focused on the impact of seasonal changes in water quality induced by precipitation on certain pathogens implicated in diarrhea, e.g. *Escherichia coli* (*E. coli*), and discovered a negative association between good water quality and high *E. coli* counts. The above analyses looked at certain summaries of spatially or temporally referenced covariates and used a generalized estimating equations approach to account for correlation over time. However, none of the analyses to date have fully exploited the spatio-temporal richness of the serial case-patterns in this data set. In addition to exploring association of diarrheal incidence with spatio-temporal covariates, the investigators

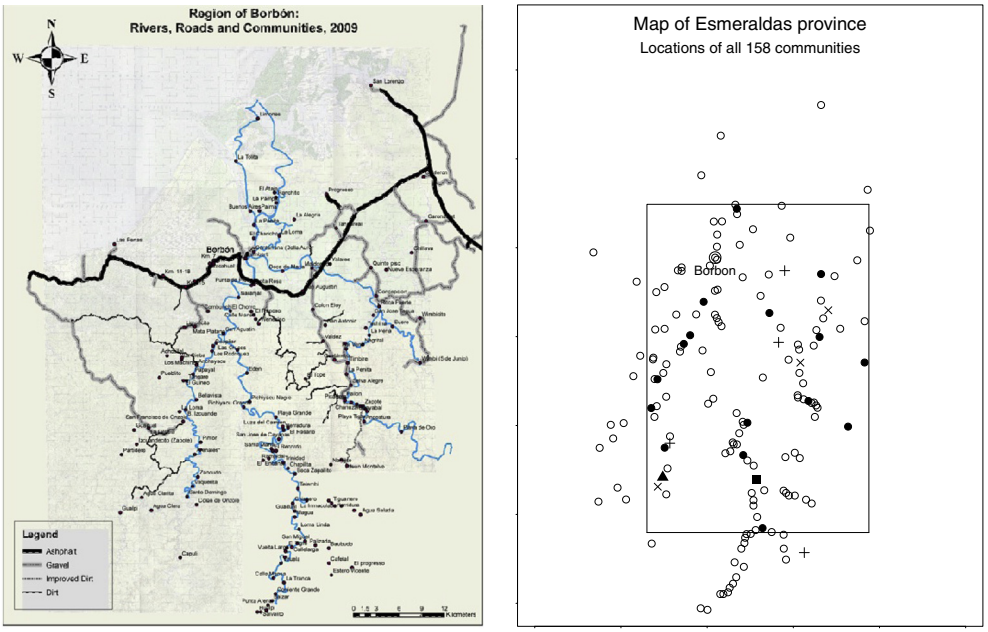

Fig. 1. The left panel displays a map of the Esmeraldes province in Ecuador. The blue lines represent three main rivers; the Cayapas, the Santiago, and the Onzole. Each point on the map indicates the location of a community. The right panel shows whether a community was sampled in the ECODESS study. Solid circles •, crosses ×, triangle ▲ and square ■ represent 21 original sampled communities whereas open circle ○ and + represent remaining unsampled communities at cycles 1–7. Here × and ▲ indicate communities used for spatial validation, × and ■ indicate communities used for temporal validation at cycle 8 and + indicates communities sampled at cycle 8, but not at cycle 1–7, used for spatio-temporal validation (see Section 2 for details). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)