Environmental Research 136 (2015) 373–380

Contents lists available at ScienceDirect

Environmental Research

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

Variability in results from negative binomial models for lyme disease measured at different spatial scales ☆

Phoebe Tran [a,*], Lance Waller [b]

[a] Department of Epidemiology, Harvard School of Public Health, Boston, MA, US
[b] Department of Bioinformatics and Biostatistics, Rollins School of Public Health, Atlanta, GA, US

ARTICLE INFO

Article history:
Received 17 June 2014
Received in revised form
19 August 2014
Accepted 21 August 2014

Keywords:
Epidemiology
Risk analysis
Lyme disease
spatial scales

ABSTRACT

Lyme disease has been the subject of many studies due to increasing incidence rates year after year and the severe complications that can arise in later stages of the disease. Negative binomial models have been used to model Lyme disease in the past with some success. However, there has been little focus on the reliability and consistency of these models when they are used to study Lyme disease at multiple spatial scales. This study seeks to explore how sensitive/consistent negative binomial models are when they are used to study Lyme disease at different spatial scales (at the regional and sub-regional levels). The study area includes the thirteen states in the Northeastern United States with the highest Lyme disease incidence during the 2002–2006 period. Lyme disease incidence at county level for the period of 2002–2006 was linked with several previously identified key landscape and climatic variables in a negative binomial regression model for the Northeastern region and two smaller sub-regions (the New England sub-region and the Mid-Atlantic sub-region). This study found that negative binomial models, indeed, were sensitive/inconsistent when used at different spatial scales. We discuss various plausible explanations for such behavior of negative binomial models. Further investigation of the inconsistency and sensitivity of negative binomial models when used at different spatial scales is important for not only future Lyme disease studies and Lyme disease risk assessment/management but any study that requires use of this model type in a spatial context.

1. Introduction

Since Lyme disease's discovery in Connecticut in the 1970s, it has been the subject of many studies due to the severity of complications that can arise in the later stages of the disease and its status as the most frequently reported vector borne illness in the United States (Bown, 2003; Feder et al., 2007). Although it has been noted that Lyme disease can potentially occur anywhere there is an intersection of Lyme disease causing ticks (*Ioxedes scapularis*), reservoir hosts, sources of blood meal, and suitable climate conditions, the highest reported incidences of Lyme disease have historically been in the Northeast, the North Central states, and the West Coast (Bown, 2003; Diuk-Wasser, 2006). As a result, most of the literature on Lyme disease has focused on these three regions.

While ticks can become infected anytime during their larval, nymph, and adult stages from any competent reservoir host, they mainly pass on the disease to humans during the nymph and adult stage of their life cycle (CDC 2014; Killilea, 2008). Transmission of *Borrelia burgdoferi*, the bacterium that causes Lyme disease, from ticks to humans occurs most frequently during late spring, early summer, and fall when nymph and adult ticks are most active (Bown, 2003). Thus, the life cycle of ticks requires animals that are infected with *Borrelia burgdoferi*, ticks that can pass on the bacteria, animal hosts that can serve as a source of blood meal for the ticks, and suitable climate conditions for the survival of ticks and their hosts (Killilea, 2008).

As the two-year life cycle of ticks is strongly influenced by various natural factors, climatic conditions, forest fragmentation, abundance of acorns, and vector hosts populations, these have all been cited in previous studies as variables that potentially have a causative relationship with Lyme disease cases (Killilea, 2008; Schauber, 2005). In order to better observe and understand the correlative relationship these variables have with Lyme disease, researchers have frequently come to rely on negative binomial models (e.g., Bouchard, 2013; Diuk-Wasser, 2012; Finch, 2014). Finch et al. utilized negative binomial regressions to determine the

☆**Funding source**: A grant from the Emory University SIRE program made it possible to purchase the technological materials (a large screen monitor and an external hard drive) that were needed for this study.

* Correspondence to: 258 English Station Road, Knoxville, TN 37934, USA.
E-mail address: pmt585@mail.harvard.edu (P. Tran).

http://dx.doi.org/10.1016/j.envres.2014.08.041
0013-9351/

relationship between landscape metrics and the density of nymphs in their study (Finch, 2014). This was done in an effort to quantify the influence that the density of infected tick nymphs and behavioral risk factors of exposure have on Lyme disease in Block Island, Rhode Island (Finch, 2014). Diuk-Wasser (2012) employed a zero inflated negative binomial model on environmental variables to approximate the density of deer tick nymphs infected with Lyme disease. The results of their negative binomial model allowed them to construct a predictive map that highlighted areas of high and low Lyme disease risk in the eastern United States (Diuk-Wasser, 2012). Zero inflated negative binomial models were also used by Bouchard et al. to better understand the factors that affected the abundance of deer tick populations in southeastern Canada (Bouchard, 2013). Using the results of their negative binomial models, the researchers of the study were able to determine that community and Lyme disease host biodiversity have some inhibitory influence on the ability of Lyme disease causing ticks to effectively pass on the disease (Bouchard, 2013).

Most of the studies on Lyme disease that use negative binomial models, mentioned earlier, only concentrate on Lyme disease at one scale. (Bouchard, 2013; Diuk-Wasser, 2012; Finch, 2014). For instance, Diuk-Wasser et al. studied Lyme disease at the local scale on Block Island, Rhode Island, while Finch et al., studied Lyme disease at the regional scale in southeastern Canada, and Bouchard et al. at the semi-national scale using 304 sites in 37 states that were east of the 100th meridian (Bouchard, 2013; Finch, 2014; Diuk-Wasser et al., 2014). However, as more and more Lyme disease studies begin to focus on Lyme disease at different spatial scales (local, sub-regional, regional, and national scale), it is important to assess the sensitivity of negative binomial models.

Currently, there are few to no studies which explore the sensitivity of negative binomial models when they are used to study Lyme disease at multiple spatial scales. In this context, our objective is to see how sensitive the negative binomial models are that are used to study the level of impact that climate factors and landscape fragmentation indicators (identified in a previous study) have on Lyme disease (Tran and Waller, 2014). If the negative binomial models prove to be too sensitive or inconsistent at multiple spatial scales, then a second objective would be to understand what might have caused these inconsistencies. In order to achieve these goals, we designed a study whose (1) study area was large enough (i.e., at a sub-regional and a regional scale) to capture the variations of key independent factors (e.g., climate, landscape fragmentation) but (2) the spatial unit was still fine enough (i.e., at local level) to capture the distinctiveness of environmental variables.

Additional motivations to conduct this study include the desire to take beginning steps towards developing more accurate and

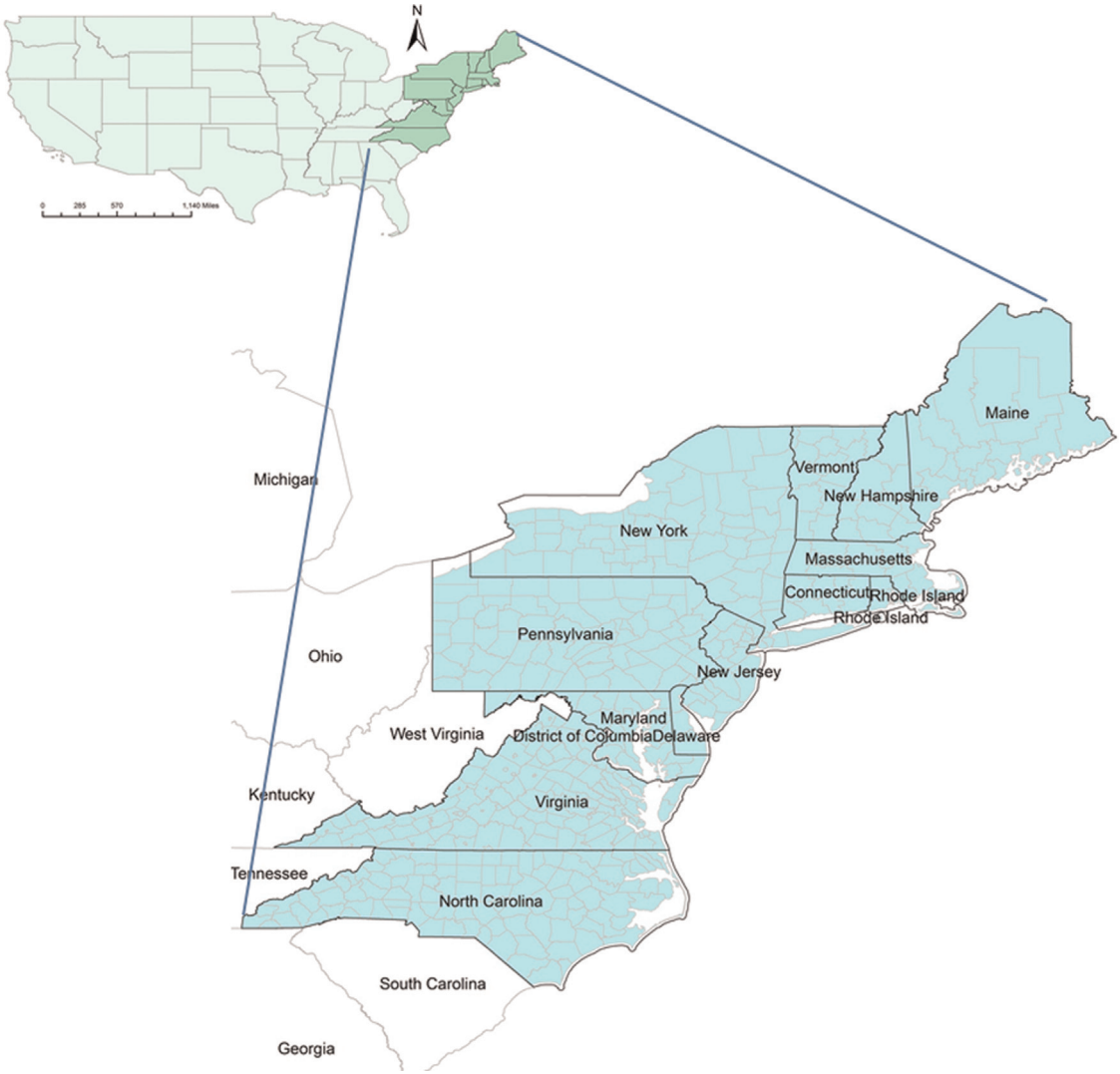

Fig. 1. Map of the study area.