Journal of Environmental Management 192 (2017) 124–133

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Application of landscape epidemiology to assess potential public health risk due to poor sanitation

Kaveh Deilami, John F. Hayes, James McGree, Ashantha Goonetilleke*

Science and Engineering Faculty, Queensland University of Technology (QUT), GPO Box 2434, Brisbane, Queensland 4001, Australia

A R T I C L E I N F O

Article history:
Received 15 November 2016
Received in revised form
19 January 2017
Accepted 22 January 2017

Keywords:
Disease risk
Geographic Information System
Hotspot analysis
Landscape epidemiology
Spatial epidemiology
Water pollution

A B S T R A C T

Clear identification of areas vulnerable to waterborne diseases is essential for protecting community health. This is particularly important in developing countries where unsafe disposal of domestic wastewater and limited potable water supply pose potential public health risks. However, data paucity can be a compounding issue. Under these circumstances, landscape epidemiology can be applied as a resource efficient approach for mapping potential disease risk areas associated with poor sanitation. However, in order to realise the full potential offered by this approach, an in-depth understanding of the impact of different classes of an explanatory variable on a target disease and the validity of hotspot analysis using limited datasets is needed. Accordingly, this research study focused on typhoid and diarrhoea incidence with respect to different classes of elevation, flood inundation, land use, soil permeability, population density and rainfall as explanatory variables. An integrated methodology consisting of hot spot analysis and Poisson regression was employed to map potential disease risk areas. The study findings confirmed the significant differences in the influence exerted by the various classes of an explanatory variable in relation to a target disease. The results also confirmed the feasibility of the hotspot analysis for identifying areas vulnerable to the target diseases using a limited dataset. The study outcomes are expected to contribute to creating an in-depth understanding of the relationship between disease prevalence and associated landscape factors for the delineation of disease risk zones in the context of data paucity.

1. Introduction

Efficient management of domestic wastewater is a primary requirement for human well-being as failure can lead to adverse public health and environmental consequences. Unfortunately, progress on urban sanitation is lagging in many developing countries (WHO, 2015). In most developing countries, reticulated sewerage systems provide only partial coverage in urban areas as the provision of conventional wastewater collection and treatment systems can be cost prohibitive. Under these circumstances, on-site treatment of wastewater such as septic tanks is an attractive alternative considering its simplicity and the relative low cost in construction and maintenance, in order to achieve the required sanitation outcomes.

Unfortunately, urban development can be in areas that are technically and environmentally unsuitable for on-site wastewater treatment due to factors such as poor soil conditions, shallow groundwater table, steep landscape and close proximity to sensitive surface water and groundwater resources. The significance stems from the fact that on-site wastewater treatment entails the discharge of partially treated sewage effluent to a surface or subsurface disposal area for the final treatment (Dawes and Goonetilleke, 2003; Khalil et al., 2004; Carroll et al., 2006). Therefore, unsatisfactory surface or subsurface conditions can result in the leaching of partially treated sewage effluent, leading to the contamination of surface and groundwater resources (Carroll and Goonetilleke, 2005). This can have serious consequences for human health and well-being in regions where reticulated water supply is not available and groundwater is used for potable purposes, particularly the shallow groundwater resources.

Considering the magnitude of the problem to be overcome in developing countries (WHO, 2015), it is imperative that limited resources available for providing proper sanitation are prudently utilised. Unfortunately, guidance is rarely available for the identification of the most vulnerable areas for the prioritisation of

* Corresponding author.
E-mail address: a.goonetilleke@qut.edu.au (A. Goonetilleke).

http://dx.doi.org/10.1016/j.jenvman.2017.01.051
0301-4797/

intervention, which poses significant challenges to decision-making. The development of technically robust methodology to identify areas vulnerable to waterborne disease risk due to poor sanitation were the key drivers for this research study. The potential occurrence of typhoid and diarrhoea, which are among the most prevalent waterborne diseases due to poor sanitation, was the key determinant (Hrudey and Hrudey, 2007).

The incidence and transmission of diseases is a function of person, place and time (Emmanuel et al., 2011). In this context, landscape epidemiology is a versatile concept to incorporate such variables to delineate the occurrence of disease. This concept explains that disease vectors, hosts and pathogens are generally tied to the landscape with ecological determinants influencing their distribution and abundance in the environment from the perspective of potential public health risk (Ostfeld et al., 2005). Past researchers have used this concept, also termed as spatial epidemiology, medical geography, and geographical epidemiology to investigate the prevalence of various diseases (Young et al., 2013; Arifin et al., 2015).

Models used in relation to landscape epidemiology investigations can be categorised into two classes, namely, spatial risk interpolation and space-time risk models (Emmanuel et al., 2011). The former (e.g. GIS-based statistical models) is employed to produce risk maps of potential disease occurrence at locations where there is no prior information about the incidence of a specific disease. The latter provides an early-warning system which can be efficient in detecting local outbreaks. The review of past landscape epidemiology studies show wide application of these models for mapping the prevalence of diseases (Real and Biek, 2007; Lambin et al., 2010; Meentemeyer et al., 2012). However, these past studies have not specifically investigated situations where data paucity is a significant issue, which require integrating both types of models to derive meaningful outcomes.

Additionally, past studies have commonly not investigated the influence of different classes of explanatory variables on study outcomes in relation to the prevalence of disease. For example, this relates to the studies undertaken by Halimi et al. (2014) for the prediction of incidence of malaria and Corner et al. (2013) in regards to the prevalence of typhoid. Accordingly, the current research study focused on the following areas: (1) to determine the contributions of various explanatory variables to landscape epidemiology study outcomes in relation to the incidence of typhoid and diarrhoea, which are common diseases associated with poor sanitation; and (2) predicting the incidence pattern of a given disease using the relationships between explanatory variables and data on the incidence of such diseases applicable to data poor areas. The study outcomes are expected to contribute to creating an in-depth understanding of the relationship between disease prevalence and associated landscape factors. More importantly, it will contribute to an improved understanding for the delineation of disease risk zones in the context of data paucity.

2. Material and methods

2.1. Study area

Semarang City, capital of Central Java Province, Indonesia was the case study area for the research project (Fig. S1 in Supplementary Information). Semarang City land area is 374 km^2 in extent with a population of over 1.6 M. Semarang is typical of most population centres in Indonesia in terms of the urban footprint and sanitation infrastructure. Overall, in Indonesia only about 72% of urban population have access to proper sanitation facilities. However, even these facilities consist of decentralised or community systems and there is no city-wide sewerage system. Consequently, the urban population is dependent on on-site wastewater treatment systems, with individual septic tanks being the most common. Due to the high population density in some areas and high water table, the most tangible consequence is poorly treated effluent being discharged to the shallow groundwater aquifer. Unfortunately, the shallow groundwater aquifer provides drinking water supply to a part of the urban population, particularly the urban poor due to the very limited pipe-borne water supply and inability to purchase bottled water. Therefore, a significant part of the urban population is subject to potentially high health risk, with typhoid and diarrhoea being among the most common diseases within the community.

2.2. Data collection

An initial desktop study identified a logical set of fundamental spatial datasets as mandatory to meet the scope of the envisaged spatial analysis. The attributes of these fundamental datasets which were considered to be causations or contributors to the location, and concentration of diseases were determined through expert-elicited knowledge underpinned by past research (Mawdsley et al., 1995; Dawes and Goonetilleke, 2003; Khalil et al., 2004; Osei et al., 2010; Seilheimer et al., 2013; Cumming et al., 2015). Integrating expert-elicited knowledge is based on the scientifically robust mechanism for incorporating expert opinion in inference (Lele and Das, 2000).

Accordingly, six variables were selected as the contributors to incidence of typhoid and diarrhoea. These variables included elevation, flood inundation, land use, soil permeability, population density and rainfall and the relevant data was obtained as spatial layers. The spatial layers then were converted to raster format. A grid cell size of 100 $\times$ 100 m was selected after evaluating a range of resolutions. It was found that the selected resolution was the most appropriate based on the envisaged analyses, computational efficiency and maintaining spatial data integrity. The spatial layers (i.e. elevation and population density) were then categorised into related classes using Jenks Natural Breaks algorithm. The categorization contributes to determine the impacts of different classes of an explanatory variable of a given disease to incidence of such disease. Fig. 1 shows the spatial layers used as explanatory variables of typhoid and diarrhoea, and associated classes. The spatial distribution of typhoid and diarrhoea are shown in Fig. 2. Note that data on the number of incidences of typhoid and diarrhoea was available only for 31 out of 172 villages in Semarang. The health authorities maintain paper-based records and it was not possible to gain access to all of the required data records.

2.3. Data analysis

2.3.1. Estimating data layer impact on diseases incidence

Each of the independent variables (i.e. spatial layers) were considered as potentially influential in determining disease potential. Therefore, in order to quantify this potential influence, the association between the data layers and disease counts for typhoid and diarrhoea needed to be estimated. This estimation was facilitated via the maximum likelihood technique for fitting Poisson regression models (McCullagh and Nelder, 1989), where each data layer in the GIS database (except rainfall) was considered as a categorical variable with groups representing classes within each layer. The initial analysis using categorical data showed rather poor relationship between each rainfall category and diseases incidence. Therefore, average rainfall amount was used as the explanatory variable. The predicted disease count for an area was taken as the exponential of the main effects linear predictor, evaluated based on the values of each GIS layer for the given area. As the study was