Science of the Total Environment 637–638 (2018) 865–870

Contents lists available at ScienceDirect

Science of the Total Environment

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

Shiga toxigenic *Escherichia coli* incidence is related to small area variation in cattle density in a region in Ireland

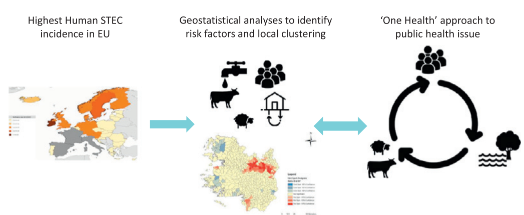
C. Brehony [a,*], J. Cullinan [b], M. Cormican [a,c], D. Morris [a]

[a] Antimicrobial Resistance and Microbial Ecology Group, School of Medicine, National University of Ireland, Galway, Ireland
[b] School of Business & Economics, National University of Ireland, Galway, Ireland
[c] Department of Medical Microbiology, University Hospital Galway, Galway, Ireland

HIGHLIGHTS

- GIS applied to infectious disease incidence in country with highest EU rates
- Cattle density associated with human STEC incidence in the West of Ireland
- Significant local clustering of STEC incidence in the region
- GIS can aid targeting public health resources for reduction of infectious disease.
- "One-health" approach necessary for public health policy

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 12 March 2018
Received in revised form 3 May 2018
Accepted 3 May 2018
Available online xxxx

Editor: D. Barcelo

Keywords:
STEC
GIS
Public-health
Clustering
Risk-factors
Zoonosis

ABSTRACT

Shiga toxigenic *Escherichia coli* (STEC) are pathogenic *E. coli* that cause infectious diarrhoea. In some cases infection may be complicated by renal failure and death. The incidence of human infection with STEC in Ireland is the highest in Europe. The objective of the study was to examine the spatial incidence of human STEC infection in a region of Ireland with significantly higher rates of STEC incidence than the national average and to identify possible risk factors of STEC incidence at area level. Anonymised laboratory records (n = 379) from 2009 to 2015 were obtained from laboratories serving three counties in the West of Ireland. Data included location and sample date. Population and electoral division (ED) data were obtained from the Irish 2011 Census of Population. STEC incidence was calculated for each ED (n = 498) and used to map hotspots/coldspots using the Getis-Ord Gi* spatial statistic and significant spatial clustering using the Anselin's Local Moran's I statistic. Multivariable regression analysis was used to consider the importance of a number of potential predictors of STEC incidence. Incidence rates for the seven-year period ranged from 0 to 10.9 cases per 1000. A number of areas with significant local clustering of STEC incidence as well as variation in the spatial distribution of the two main serogroups associated with disease in the region *i.e.* O26 and O157 were identified. Cattle density was found to be a statistically significant predictor of STEC in the region. GIS analysis of routine data indicates that cattle density is associated STEC infection in this high incidence region. This finding points to the importance of agricultural practices for human health and the importance of a "one-health" approach to public policy in relation to agriculture, health and environment.

1. Introduction

Shiga toxigenic *Escherichia coli* (STEC) also known as verocytotoxigenic *E. coli* (VTEC) are pathogenic *E. coli* that are associated with a spectrum of illness. After exposure some people remain asymptomatic whereas others

* Corresponding author.
E-mail address: carina.brehony@nuigalway.ie (C. Brehony).

https://doi.org/10.1016/j.scitotenv.2018.05.038
0048-9697/© 2018 Published by Elsevier B.V.

develop self-limiting gastroenteritis or bloody diarrhoea (haemorrhagic colitis). Haemolytic uraemic syndrome (HUS), characterized by acute renal failure, occurs in 3–7% of cases, with those most at risk being those aged <5 years of age (HPSC, 2016). The natural host for STEC is ruminants. Some animals excrete the microorganism in very high numbers (termed 'super-shedders') in their faeces (Murphy et al., 2016; Chase-Topping et al., 2008). The risk to humans is posed from contact with animal faecal matter, through direct contact with the animals or indirectly through the contaminated environment *e.g.* recreational or drinking water, or contaminated foodstuffs *e.g.* raw milk, raw meat (Kintz et al., 2017). Managing the risk to humans from STEC is particularly challenging as it has a very low infectious dose (<100 bacterial cells) and it can survive in the environment for extended periods of time (Money et al., 2010).

The incidence of human infection with STEC in Ireland has been the highest in Europe since 2008; it was ten times the EU average in 2016 with 15.6 cases per 100,000 inhabitants in Ireland *versus* an average EU incidence of 1.77/100,000 (ECDC, 2017). Within Ireland there is significant regional variation in the incidence of STEC infection. Among the high incidence areas is a region of three counties (Galway, Mayo and Roscommon) in the west of Ireland. In this region the crude incidence in 2016 was 26 per 100,000 inhabitants compared to the national average of 18 per 100,000 inhabitants (HPSC, 2016). By contrast the incidence in a region in the east of Ireland comprised of counties Dublin, Kildare and Wicklow was 7.7 per 100,000 inhabitants.

In Ireland, the most commonly recognised modes of STEC transmission are direct animal contact, person-to-person and waterborne (HPSC, 2016; Garvey et al., 2016). However, for a significant number (13%) of cases each year, the mode of transmission is reported as 'unknown' or 'unspecified' (HPSC, 2016). Small group water supplies, untreated private wells, livestock density and domestic wastewater treatment have all been implicated in STEC transmission and in STEC outbreaks (HPSC, 2016; O'Sullivan et al., 2008; ÓhAiseadha et al., 2017).

The aim of this project was to examine the spatial incidence of STEC infection in a region of Ireland with higher incidence than the national average and to model the relationship between STEC incidence and a range of potential predictors at a geographical level. Diagnostic testing for STEC in the region is concentrated primarily in one laboratory which facilitated data access.

2. Methods

Ethical approval for this study was granted by the National University of Ireland, Galway Research Ethics Committee. Anonymised laboratory data for confirmed STEC cases from 2009 to 2015 inclusive were obtained from diagnostic laboratories serving the three counties Galway, Mayo and Roscommon. Duplicates and possible linked confirmed STEC cases were removed leaving 'primary cases' ($n = 379$). Possible linked STEC cases were defined as those that were from the same electoral division (ED) with the same serogroup and within one month of a 'primary case'. Metadata recorded for each case included location geocoded at ED level and clinical laboratory sample receipt date. There were 498 EDs in the region studied. Population and ED data including number of households, land surface area (km^2), household water source and household wastewater treatment, were obtained from the 2011 Census of Population for Ireland from the Central Statistics Office (CSO, 2011). ED population ranged from 83 to 14,384 individuals and total land surface area ranged from 0.56 km^2 to 162.3 km^2. Population data was used to calculate the cumulative incidence rate (per 1000 inhabitants) of confirmed STEC cases for each ED for the study period. Data on cattle and sheep numbers were obtained from the 2010 Census of Agriculture (CSO, 2012).

The calculated STEC incidence rate was used to investigate geographic clustering within the region using 'hotspot' analysis and 'cluster' analysis. The hotspot analysis involved calculating the Getis-Ord Gi* (Getis and Ord, 1992) statistic for STEC incidence in each ED with the resulting z-scores indicating where EDs with either high or low values cluster spatially. For statistically significant positive/negative z-scores, the larger/smaller the z-score is, the more intense the clustering of high/low values *i.e.* a hotspot/coldspot. In order to be classified as a statistically significant hotspot, for example, an ED will have a high z-score and be surrounded by other EDs with high z-scores. Cluster analysis, on the other hand, uses Anselin's Local Moran's I test statistic (Anselin, 1995). Again the approach involves the calculation of a test statistic (*i.e.* Local Moran's I value) along with a z-score and a pseudo p-value, indicating the cluster type (or outlier) for each statistically significant ED. The cluster/outlier type distinguishes between a statistically significant cluster of high values (HH), cluster of low values (LL), outlier in which a high value is surrounded primarily by low values (HL), and outlier in which a low value is surrounded primarily by high values (LH). Statistical significance was set at 95%.

Finally, to consider the relationship between STEC incidence by ED and a range of potential predictors, correlation and multivariable regression analyses were undertaken. First, Pearson correlations between ED-level STEC incidence and the various potential predictors were estimated to give an indication of the strength of linear association between variables. Associated significance levels of each correlation coefficient were at the 95% level. In the multivariable analysis, STEC incidence rate in an ED was the dependent variable and a range of potential predictors were considered (Table 1), based on data availability and previously published research (CSO, 2011; CSO, 2012; Haase et al., 2014; Teljeur and Kelly, 2008). For the multivariable regression analysis, we estimated a stepwise (backward selection) model, applying a 5% significance level for removal. This meant that the variable with the greatest *p*-value was removed, one at a time, until all remaining variables had a p-value less than the 5% threshold. We also tested the robustness of our findings by applying a 10% significance level for removal, as well as by estimating a number of backward stepwise models with 5% and 10% significance levels for removal and addition. Overall, our main findings were robust across these different approaches. Moreover, we also considered models that accounted for spatial dependence in both the dependent and independent variables and found this did not alter our key conclusions.

3. Results

3.1. STEC cases and incidence

The overall seven-year incidence ranged from 0 to 10.9 cases per 1000 inhabitants across the region (Fig. 1) with a mean of 0.95/1000 – equivalent to 94.57/100,000 and a per year average of 13.51/100,000. There

Table 1
Variable definitions and sample descriptive statistics.
(Source: analysis of hospital records, Census of Population 2011 data (CSO, 2011), Census of Agriculture 2010 data (CSO, 2012), data from Haase et al. (2014) (Haase et al., 2014) and Teljeur and Kelly (2008) (Teljeur and Kelly, 2008).)

Variable name	Variable description	Mean (SD) or %
Dependent variable		
STEC rate	STEC incidence rate (cases per 1000 population)	0.946 (1.568)
Potential predictors		
Cattle density	Number of cattle per km^2	74.313 (46.031)
Sheep density	Number of sheep per km^2	90.328 (91.853)
Private water Source	Percentage of households with a private water source	0.320 (0.261)
SEPTIC tank density	Percentage of households with a septic tank	0.741 (0.262)
Deprivation	Relative deprivation (Trutz Haase Index)	−6.730 (3.450)
Population density	Number of persons per km^2	138.829 (516.467)
Rurality	=Rural ED	90.56%
	=Urban ED	9.44%

Note: all variables are calculated at ED level.