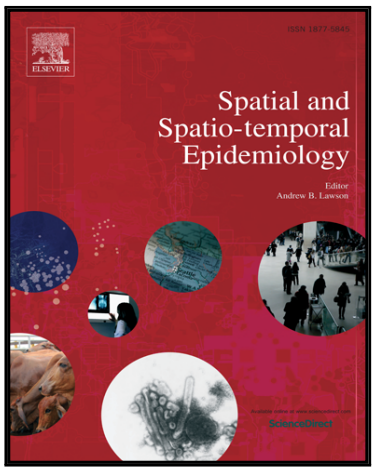
Accepted Manuscript



A Bayesian spatio-temporal framework to identify outbreaks and examine environmental and social risk factors for infectious diseases monitored by routine surveillance

Aparna Lal , Jonathan Marshall , Jackie Benschop ,
Aleisha Brock , Simon Hales , Michael G Baker , Nigel P French

PII: S1877-5845(16)30071-5
DOI: 10.1016/j.sste.2017.10.004
Reference: SSTE 229

To appear in: *Spatial and Spatio-temporal Epidemiology*

Received date: 24 August 2016
Revised date: 21 June 2017
Accepted date: 17 October 2017

Please cite this article as: Aparna Lal , Jonathan Marshall , Jackie Benschop , Aleisha Brock , Simon Hales , Michael G Baker , Nigel P French , A Bayesian spatio-temporal framework to identify outbreaks and examine environmental and social risk factors for infectious diseases monitored by routine surveillance, *Spatial and Spatio-temporal Epidemiology* (2017), doi: 10.1016/j.sste.2017.10.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Highlights

- A Bayesian spatio-temporal framework proposed to identify outbreaks and examine risk factors from routine surveillance data detected previously unidentified disease clusters and risk factors associated with reported cryptosporidiosis and giardiasis

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

A Bayesian spatio-temporal framework to identify outbreaks and examine environmental and social risk factors for infectious diseases monitored by routine surveillance

Aparna Lal[1*] aparna.lal@anu.edu.au, Jonathan Marshall[2], Jackie Benschop[2], Aleisha Brock[3], Simon Hales[4], Michael G Baker[4], Nigel P French[2]

[*]Corresponding author

[a]National Centre for Epidemiology and Population Health, Research School of Population Health, Australian National University, Canberra, Australian Capital Territory, Australia, Email: . Telephone: 0061-02-61252309.

[b]mEpiLab, Infectious Disease Research Centre, Hopkirk Research Institute, Massey University, Palmerston North, New Zealand.

[c]School of Nursing & Midwifery, Division of Health Sciences, University of South Australia, Adelaide, South Australia, Australia.

[d]Department of Public Health, University of Otago School of Medical and Health Sciences, Wellington, New Zealand.

Abstract

Spatio-temporal disease patterns can provide clues to etiological pathways, but can be complex to model. Using a flexible Bayesian hierarchical framework, we identify previously undetected space-time clusters and environmental and socio-demographic risk factors for reported giardiasis and cryptosporidiosis at the New Zealand small area level. For giardiasis, there was no seasonal pattern in outbreak probability and an inverse association with density of dairy cattle ($\hat{\beta}_1$= -0.09, Incidence Risk Ratio (IRR) 0.90 (95% CI 0.84, 0.97) per 1 log increase in cattle/km^2). In dairy farming areas, cryptosporidiosis outbreaks were observed in spring. Reported cryptosporidiosis was positively associated with dairy cattle density: $\hat{\beta}_1$= 0.12, IRR 1.13 (95% CI 1.05, 1.21) per 1 log increase in cattle/km^2 and inversely associated with weekly average temperature: $\hat{\beta}_1$=-0.07, IRR 0.92 (95% CI 0.87, 0.98) per 4°C increase. This framework can be generalized to determine the potential drivers of sporadic cases and latent outbreaks of infectious diseases of public health importance.

Keywords

Spatial, Temporal, Surveillance, Environmental, Outbreaks, Infections

ACCEPTED MANUSCRIPT

Introduction

Global environmental changes, especially climate change and human exploitation of productive ecosystems (1-3) have important implications for infectious disease risk (4, 5). For human pathogens with environmental reservoirs, such as livestock, understanding geographical and seasonal variability in risk factors can help to identify high risk locations and time-periods and predict disease incidence under scenarios of global environmental and social change (6, 7). New modelling tools can help to understand how these environmental and socio-demographic factors interact to drive disease patterns and translate this understanding to improve decision-making (8).

Cryptosporidiosis and giardiasis are infectious gastrointestinal diseases caused by the parasites *Cryptosporidium* and *Giardia* (9). Cryptosporidiosis and giardiasis are recognized by the World Health Organization as infections of global importance (10), with high disease rates among children, the elderly, socio-economically disadvantaged and immune-suppressed people (11). The parasites are primarily spread through contaminated drinking or recreational water; however, infection in humans may arise through contaminated food, contact with animals, especially livestock or infected individuals (9, 10). In New Zealand, previous research on human cryptosporidiosis and giardiasis has suggested that local weather variability, socio-economic status and degree of urbanization are important determinants of spatial disease patterns (12, 13), while seasonal animal pathogen load and host behaviour influence temporal disease patterns (14-16).

As rates of reported cryptosporidiosis and giardiasis in New Zealand continue to be higher than comparable rates in the United States (17) and Australia (18), identifying the environmental and socio-demographic exposures can help develop disease control priorities in high risk areas.

We used a Bayesian hierarchical modelling framework to identify space-time clusters of cryptosporidiosis and giardiasis in New Zealand and calculated an outbreak probability for each identified cluster. We then used a Bayesian spatio-temporal process model to identify the small area level environmental, socio-economic and demographic factors associated with the risk of reported disease.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Methods

Notification data

All notified cases of cryptosporidiosis and giardiasis during 1997-2008 in New Zealand were obtained from the National Notifiable Disease Surveillance system. The reason for choosing this time period was that this was a period of rapid change in livestock farming across New Zealand, with an increase in dairy cattle numbers, decrease in number of farms and increase in stock density (Figure S1 Supplementary Material). Further, no major changes were made to the surveillance of these notifiable diseases from 1997–2007; direct laboratory notification began in 2008. For notifications, cases were defined as a clinical illness with appropriate laboratory confirmation. Based on home address, each notified illness was assigned a 2006 Census Area Unit (CAU) code. CAUs have a population of approximately 3000-5000 people (Figure 1).

Nearly all cryptosporidiosis cases (9848 out of 9900, 99.5%) and giardiasis cases (19470 out of 19553, 99.6%) were geocoded to a CAU.

Analysis of risk factors

For this analysis, notification data were restricted to the period 2000-2007 to match data on livestock densities and ensure that direct laboratory notification introduced in 2008 did not influence patterns. We extracted laboratory confirmed cases of cryptosporidiosis (N=8688) and giardiasis (N=16930) along with the following case information: report date, age, prioritised ethnicity, and CAU code of residence. To avoid a "weekday effect" on notifications, cases were aggregated weekly where each week started on a Wednesday and ended on a Tuesday for every week in the eight year period.

Population at risk

We obtained population estimates for census years 1996, 2001 and 2006 (19). These data were linearly interpolated to produce population estimates across the study period. For the year 2007, the population estimates were extrapolated from the trend. The total population as well as the population by age group was provided where population estimates at 30 June 2000 were based on 2001 CAU boundaries, while population estimates from 2001 onwards were based on 2006 CAU boundaries. Using a geographic concordance file, the 2001 CAU boundaries were matched to the corresponding 2006 CAUs. Population was used as an offset to ensure that the case numbers were adjusted for the population at risk and population density was used as a

ACCEPTED MANUSCRIPT

covariate in the analysis to assess non-linear effects of density. Population density was calculated by dividing the total population number in each CAU by the area in each CAU in ArcGIS v10.1(20).

Urban/Rural residence

For each CAU, Statistics New Zealand categories (21) 'main urban areas', 'satellite urban areas' and 'independent urban areas' were classified as urban (reference category) whereas the categories 'rural areas with high urban influence', 'rural areas with moderate urban influence', 'rural areas with low urban influence' and 'highly rural/remote areas' were classified as rural.

New Zealand Deprivation Index

We estimated deprivation by CAU for study years by linear interpolation of the New Zealand Index of Deprivation (NZDep) for 2001 and 2006 (22). Deprivation levels were grouped into terciles, with levels 1-3 representing affluent CAUs (Level 1) (reference category), 4-7 representing medium (Level 2) and 8-10 representing the least affluent CAUs (Level 3). These were fitted as a categorical variable.

Age and Ethnicity estimates

Ethnicity estimates for the years 2001 and 2006 by CAU were provided by Statistics New Zealand and were based on 2006 CAU boundaries. For each major ethnic group, the percentage of the population number in each ethnic group to the total CAU population number of all ethnic groups in the CAU was calculated. The estimates for every year were calculated by linear interpolation of census years. For the year 2007, the population estimates were extrapolated from the trend. Ethnicity was based on Level 1 prioritised ethnicity which divides the population into European, Māori, Pacific Peoples, Asian, Middle Eastern/Latin American/African (MELAA) and Other. We estimated the proportion of the CAU population in each of three age categories (0-4; 5-64 and ≥65 years) based on interpolation of census data, as in a previous study (23).

Drinking water quality

Annual drinking water quality grading was supplied by the Institute of Environmental Science and Research (ESR) Water Programme for the years 2000-2007. ESR used both the distribution zone code and protozoa compliance to construct a scoring system for drinking water quality. A distribution zone is defined as "all or part of a reticulated supply for which the water is expected to be of consistent quality throughout" (24). Protozoa compliance refers to compliance at the treatment plant and is based on "monitoring the effectiveness

ACCEPTED MANUSCRIPT

of the treatment processes used to remove or disinfect *Cryptosporidium*" (24). Following Brock (2011), a grid score of 0 denoted good drinking water quality (complied); a grid score of 1 denoted intermediate drinking water quality (inadequately monitored); a grid score of 2 represented poor drinking water quality (non-compliant and either not monitored or contained *E.coli*); and a grid score of 3 indicated the drinking water quality was unknown. The method of assigning drinking water quality to CAUs is detailed in in the Supplementary Material (Methods Section 1)

Livestock density

Numbers of dairy cattle, beef cattle, sheep, pigs, poultry and deer for each farm in New Zealand were obtained from the Agribase^{TM} database for each alternate year from 2000-2008. Livestock densities were calculated by dividing the number of animals in each CAU by the total land area in each CAU. Data for the missing years were interpolated. We included the logarithm of livestock density in the model as these data are skewed to the right (Figure S2). The method of assigning livestock numbers from each farm to a CAU is detailed in the Supplementary Material (Methods Section 2).

Climate data

Weekly time series of average temperature and rainfall for CAUs based on interpolated observations from land stations were provided by The National Institute of Water and Atmospheric Research (NIWA).These data were standardised to have mean 0 and standard deviation 1.

Data analysis

Spatial and temporal patterns in disease risk

To assess spatial and temporal patterns in disease risk, we used the Bayesian hierarchical model described by Spencer et al 2011 (25).

Bayesian model definition

The number of cases in CAU *i* in week *t* is denoted by Y_{it} and assume that Y_{it} ~ Poisson $(n_{it}\lambda_{it})$ where the offset n_{it} is the population in CAU *i* in week t and $\lambda_{it}$is the risk associated with CAU *i* in week *t*. The log of the risk was decomposed into four components: an intercept α, a purely temporal component R_t; a purely spatial component U_i; and the spatio-temporal component W_{it}.

Equation 1

$$\log\left(\lambda_{it}\right) = \alpha + R_t + U_i + W_{it}$$

The temporal and spatial terms R and U are modelled via structural priors as random effects. We assume the risk in time *t + 1* is a linear extrapolation of risk at times *t* and *t-1*

$$R_{t+1} \sim Normal\left(2R_t - 2R_{t-1}, \sigma^2{}_R\right)$$

Where sigma is the variance component of the normal distribution.

Risk in spatial unit *j* is modelled as the average of risk in neighbouring area units

$$U_i \sim Normal\left(\sum_j \frac{U_j}{n_i}, \frac{\sigma_U^2}{n_i}\right)$$

Where *j* is taken over the n_i area units neighbouring area unit *i*.

Note that R, U and W are treated as random effects, and thus account for overdispersion (Equation 1). In particular, the spatio-temporal component W_{it} term accounts for increased (or decreased) risk in a census area unit during each week (26) and is designed to capture short term localised periods of increased risk that are typical of outbreaks (described in the next section).

Outbreak analysis

For the analysis of outbreaks and risk factors, the small island CAUs, those classified as harbours, inlets and oceanic regions and unpopulated CAUs were excluded. CAUs that had missing data for exposures of interest (such as urban/rural status, Deprivation Index) were also excluded. This resulted in a total of 1778/1927 (92 %) of CAUs (mainland populations with all the covariates of interest) being included in the final analysis.

The model developed using Equation 2 was applied to each disease dataset to detect outbreaks clustered in time and space. In order to do this, Territorial Authorities (TAs, n=72) were used to group CAUs (n=1778) into regions (Figure 1). This reduced the number of parameters that needed to be estimated in the models and improved our ability to identify "outbreaks", as they were based on several observations rather than a single datum. Given equation 1, we set

ACCEPTED MANUSCRIPT

Equation 2

$$W_{i,t} = \beta_{r(i)} X_{r(i),t}$$

Where $r(i)$is the index of the region containing CAU i, $X_{r,t}$ is the outbreak indicator and β_r reflects the size of the increase in risk for outbreaks in region r. The posterior proportion where $X_{r,t}$ = 1 thus gives the probability of an anomalous event. For detailed assessment of spatial and temporal patterns, one predominantly urban TA (Auckland) and one predominantly rural TA (Clutha) were chosen (Figure 1). Full details of this model, including priors and Monte Carlo Markov Chain fitting scheme have been published (25).

Analysis of risk factors

The historical relationship between climatic, land use and social factors with cryptosporidiosis and giardiasis incidence was analysed using an extension of the outbreak model (Equation 3). Seasonality was assumed to be absorbed by the flexible temporal component, R_t and the spatio-temporal term W_{it} being given by Equation 3.

$$W_{it} = \beta_1 Z_{1it} + \beta_2 Z_{2it} + \ldots + \beta_k Z_{kit}$$

Where Z_{1it}, ...,Z_{kit} are explanatory covariates, such as weather, demographic and environmental variables, and $\beta_1,\ldots,\beta_k$ are coefficients to be estimated. Uninformative normal priors were specified for β_k.

To attain convergence, 4 separate chains of 100,000 iterations sampling every 50th iteration after an initial burn-in of 5000 iterations were run for each model. Having adjusted for the effect of other covariates, the change in incidence risk (IR) of each disease was calculated using Equation 4.

Equation 4

$$\Delta\lambda_{it} = e^{\hat{\beta}_k}$$

where $\Delta\lambda_{it}$ represents the change in rate of a notified case of protozoan disease in grid i during week t; and $\hat{\beta}_k$ represents the posterior mean coefficient of each covariate Z_{kit}. Model coefficients are related to incidence risk as follows: a change in temperature or rainfall equivalent to one standard deviation; for density variables, unit change on a log scale; for proportions (age, ethnicity), a change of one percent; for categorical variables (rurality, drinking water, deprivation) change compared to the baseline.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Spatial and temporal patterns in disease risk

Figure 2 shows the relative risks estimated by the spatial component of the model for cryptosporidiosis and giardiasis across the whole country. The relative risk is interpreted as the risk of disease in each CAU as compared to the average risk across all units. Therefore, a value greater than 1.0 implies a higher than average risk, whereas a value less than 1.0 implies a lower than average risk.

Figure 3 shows the relative risks estimated by the spatial component of the model for cryptosporidiosis and giardiasis in the urban Auckland region. The relative risk of giardiasis was spatially heterogeneous, with many of the high risk areas appearing along the coast (Figure 3A). In contrast, the relative risk of cryptosporidiosis notification in urban areas of Auckland was more uniform and was less than the average risk expected across all CAUs (Figure 3B).

Figures 3C and 3D show the relative risks estimated by the spatial component of the model for giardiasis and cryptosporidiosis in the rural Clutha region. In the rural Clutha District, the relative risk for giardiasis notifications was below the average estimated for all CAUs across the entire area (Figure 3C). Conversely, for cryptosporidiosis, the risk of notification was highest in the Clutha region as compared to the average expected risk (Figure 3D).

Figure S3 shows the temporal trend of giardiasis and cryptosporidiosis notifications across the study period. For giardiasis, a decreasing trend until early 2000 was followed by a fairly consistent number of reported cases with seasonal fluctuations, but no distinct peaks. The trend for the average number of cryptosporidiosis cases remained fairly similar across the entire time period with clear dominant peaks in spring with two apparent autumn peaks in early years.

The posterior outbreak probabilities for the Auckland and Clutha Districts for cryptosporidiosis and giardiasis are shown in Figure 4. For giardiasis, in Auckland, there were no clear patterns in the outbreak probabilities estimated (Figure 4A). In the Clutha District for giardiasis there was no evident pattern in the estimated posterior outbreak probability (Figure 4C). For cryptosporidiosis, in Auckland, there were a number of distinct periods of increased outbreak probability, for example during the first half of 2001 and the first half of 2007 (Figure 4B). For cryptosporidiosis in the Clutha District, distinct spring peaks in the outbreak probability were

ACCEPTED MANUSCRIPT

observed in most years, particularly during the spring of 2000, 2001, 2002 and a smaller summer peak in 2000 was seen (Figure 4D).

Risk factors

Giardiasis

Having adjusted for the effect of other covariates, the variables associated with an increased risk of giardiasis were the percentage of the population being less than four years old ($\hat{\beta}_1$= 0.04, IRR 1.04 (95% CI 1.02, 1.06)) (Table 1). Factors inversely associated with the risk of giardiasis included the percentage identifying with Asian, Māori or Pacific Island ethnic groups, the percentage over 65 years old, log population density (people/km^2) ($\hat{\beta}_6$=-0.17, IRR 0.84 (95% CI 0.76, 0.92)) and log dairy cattle density (cows/km^2) ($\hat{\beta}_7$= -0.10, IRR 0.91 (95% CI 0.85, 0.98)) (Table 1).

Cryptosporidiosis

Having adjusted for the effect of other covariates the variables positively associated with cryptosporidiosis risk were the percentage of the population less than four years old ($\hat{\beta}_1$= 0.04, IRR 1.04 (95% CI 1.02, 1.06)) and log dairy cattle density (cows/km^2) ($\hat{\beta}_1$= 0.12, IRR 1.13 (95% CI 1.05, 1.21)) (Table 2). Factors inversely associated with the risk of cryptosporidiosis were the percentage identifying with Asian, Māori or Pacific Island ethnic groups, the percentage over 65 years old, and log population density (people/km^2) ($\hat{\beta}_8$=-0.30, IRR 0.74 (95% CI 0.69, 0.80)) (Table 2).Figure S4 shows the increase in the estimated IRR of cryptosporidiosis relative to a baseline of 1 cow/ km^2.

Discussion

We identified previously undetected clusters and estimated the independent effects of key physical and climate variables, environmental exposures, and demographic and socio-economic factors on the risk of infectious diseases monitored through routine surveillance. Neighbourhood environmental and social factors are associated with the risk of reported giardiasis and cryptosporidiosis in New Zealand, with distinct spatial and temporal patterns in plausible outbreaks. For both diseases, the risk of reported illness was negatively associated with the percentage of Asian, Māori, and Pacific Island ethnic groups, those over 65 years and population density and positively associated with the percentage of children up to 4 years old. Giardiasis was

ACCEPTED MANUSCRIPT

negatively associated with dairy cattle density. Cryptosporidiosis was positively associated with dairy cattle density and inversely associated with weekly temperature.

Spatial patterns in reported giardiasis suggest higher risk in urban areas, with little seasonality in probable outbreaks. Giardiasis tends to be the least seasonal of the important enteric zoonotic infections globally (27). In New Zealand, meteorological conditions such as temperature and rainfall are not significantly associated with this enteric infection (28). Results of a case-control study in Auckland suggest that changing nappies (diapers) is a significant risk factor (29). Descriptive studies showing increased risk in the age group 30-39 years (13) and in females (30), may also be related to a closer association with toddlers. Attendance at day care centres has been frequently cited as a common risk factor for giardiasis in children (31) and may be related to the higher risk that we found in urban areas. Our findings of a significantly higher risk of reported giardiasis in children up to 4 years old further support this hypothesis.

Spatial patterns in reported cryptosporidiosis suggest that for this disease, environmental sources of infection dominate. The uniformly high risk patterns in the Clutha District, an area with high dairy cattle densities, and the clear recurrent spring peaks in plausible outbreaks that is absent in urban regions imply a common environmental exposure that is likely to be related to agricultural activities such as contact with newborn livestock. This conclusion is further supported by a significant, positive association of disease risk with dairy cattle density. Cryptosporidiosis outbreaks have been reported following farm visits (32) and among veterinary students in New Zealand (33). Genotyping of human and animal isolates in New Zealand, the United Kingdom and Australia suggest that sporadic cryptosporidiosis in rural areas is primarily driven by zoonotic transmission (14, 34, 35). Seasonal patterns may also be due to the effect of ambient environmental conditions on pathogen survival (36) and transport (37) or the role of environmentally sensitive vectors such as flies (38), or rural practitioners requesting more intensive testing for *Cryptosporidium* spp.. In areas where dairy farming is predominant, the higher risk of reported cryptosporidiosis and seasonal patterns of outbreaks suggests an important role for environmental transmission.

Ethnic and socio-economic disparities tend to reduce health care access and utilisation by vulnerable populations, particularly at the primary care and laboratory diagnosis level. This effect limits the extent to which notification data may be used to infer causal relationships between disease risk and ethnic and socio-

ACCEPTED MANUSCRIPT

economic factors. Such biases may result in under representation of vulnerable groups when using data from passive surveillance systems (39). For both diseases, the percentage of Asian, Māori, and Pacific Island ethnic groups in an area were inversely associated with disease risk. This inverse association is in keeping with results from previous studies in New Zealand (12, 13) and is probably indicative of a significant bias in case ascertainment rather than an effect of ethnicity. For enteric salmonellosis, opposite trends in hospitalisations for these ethnic groups are seen, suggesting more severe outcomes for these populations (40). The finding from this study that the more socio-economically deprived areas were inversely associated with cryptosporidiosis and giardiasis risk (Table 1 and Table 2) is likely to be an artefact of passive surveillance data. Although high rates of reported cryptosporidiosis and giardiasis may partly be explained by healthcare seeking behaviour, the highest rates of hospitalisations are also seen in these age groups, suggesting they are more vulnerable to symptomatic disease.

In conclusion, using contemporary modelling methods, our study builds on previous evidence by using routinely collected surveillance data to identify localised, short-term periods of increased risk and environmental, socio-economic and demographic risk factors for infectious disease. In New Zealand, the spatial and temporal variations in the risk of parasitic diseases appear to be quite distinct. Using an innovative Bayesian approach to detect spatially-localised periods of increased disease incidence, and identify associated environmental and socio-demographic exposures, can provide important evidence to guide the development of targeted disease control measures in areas where the environmental health risks are high.

ACCEPTED MANUSCRIPT

References

1. Millennium Ecosystem Assessment. Living beyond our means: natural assets and human well-being. 2005.
2. Tilman D, Balzer C, Hill J, Befort BL. Global food demand and the sustainable intensification of agriculture. P Natl Acad Sci USA. 2011;108(50):20260-4.
3. IPCC. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge, UK and New York, NY, USA: Cambridge University Press2012.
4. Murray KA, Daszak P. Human ecology in pathogenic landscapes: two hypotheses on how land use change drives viral emergence. Curr Opin Vir. 2013;3(1):79-83.
5. Cutler SJ, Fooks AR, van der Poel WHM. Public Health Threat of New, Reemerging, and Neglected Zoonoses in the Industrialized World. Emerg Infect Dis. 2010;16(1):1-7.
6. Anyamba A, Linthicum KJ, Small JL, Collins KM, Tucker CJ, Pak EW, et al. Climate teleconnections and recent patterns of human and animal disease outbreaks. PLoS Negl Trop Dis. 2012;6(1):e1465.

ACCEPTED MANUSCRIPT

7. Jones BA, Grace D, Kock R, Alonso S, Rushton J, Said MY, et al. Zoonosis emergence linked to agricultural intensification and environmental change. Proc Natl Acad Sci USA. 2013 May 21, 2013;110(21):8399-404.
8. Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. Effects of environmental change on zoonotic disease risk: an ecological primer. Trends Parasitol. 2014;30(4):205-14.
9. Cacciò SM, Thompson RA, McLauchlin J, Smith HV. Unravelling cryptosporidium and giardia epidemiology. Trends Parasitol. 2005;21(9):430-7.
10. Savioli L, Smith H, Thompson A. *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. Trends Parasitol. 2006;22(5):203-8.
11. Torgerson PR, Macpherson CNL. The socioeconomic burden of parasitic zoonoses: Global trends. Vet Parasitol. 2011;182(1):79-95.
12. Britton E, Hales S, Venugopal K, Baker MG. The impact of climate variability and change on cryptosporidiosis and giardiasis rates in New Zealand. J Water Health. 2010;8(3):561-71.
13. Snel SJ, Baker MG, Kamalesh V, French N, Learmonth J. A tale of two parasites: the comparative epidemiology of cryptosporidiosis and giardiasis. Epidemiol Infect. [Article]. 2009 Nov;137(11):1641-50.
14. Learmonth J, Ionas G, Pita A, Cowie R. Seasonal shift in *Cryptosporidium parvum* transmission cycles in New Zealand. J Eukaryot Microbiol. 2001;Suppl:34S-5S.
15. Learmonth JJ, Ionas G, Ebbett MA, Kwan ES. Genetic characterization and transmission cycles of *Cryptosporidium* species isolated from humans in New Zealand. Appl Environ Microbiol. 2004 Jul;70(7):3973-8.
16. Al Mawly J, Grinberg A, Velathanthiri N, French N. Cross sectional study of prevalence, genetic diversity and zoonotic potential of *Cryptosporidium parvum* cycling in New Zealand dairy farms. Parasit Vectors. 2015;8(1):1-7.
17. Yoder JS, Beach MJ. Cryptosporidium surveillance and risk factors in the United States. Exp Parasitol. [Article]. 2010 Jan;124(1):31-9.
18. Lal A, Cornish LM, Fearnley E, Glass K, Kirk M. Cryptosporidiosis: A Disease of Tropical and Remote Areas in Australia. PLoS Negl Trop Dis. 2015;9(9):e0004078.
19. Historical Population Estimates. [database on the Internet]. Government of New Zealand. 2010.
20. ESRI. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute; 2011.
21. Statistics New Zealand. New Zealand: An Urban/Rural Profile Update. Wellington: Statistics New Zealand; 2012 [cited 2013 30 July]; Available from: http://www.stats.govt.nz/browse_for_stats/people_and_communities/Geographic-areas/urban-rural-profile-update.aspx.
22. Salmond C, Crampton P, Atkinson J. NZDep2006 Index of deprivation user's manual- research report: Wellington School of Medicine and Health Sciences, Department of Public Health;2006.
23. Brock A. The effect of climate variation on infectious diseases in humans in New Zealand. Palmerston North, New Zealand: Massey University; 2011.
24. Ministry of Health. Annual Report on Drinking-water Quality 2013–2014. In: Health Mo, editor. Wellington: Ministry of Health; 2015.
25. Spencer SE, Marshall J, Pirie R, Campbell D, French NP. The detection of spatially localised outbreaks in campylobacteriosis notification data. Spatial and Spatio-temporal Epidemiology. 2011;2(3):173-83.
26. Hinde J, Demétrio CG. Overdispersion: models and estimation. Computational Statistics & Data Analysis. 1998;27(2):151-70.
27. Lal A, Hales S, French N, Baker MG. Seasonality in human zoonotic enteric diseases: A systematic review. Plos One [serial on the Internet]. 2012; 7(4): Available from: http://dx.doi.org/10.1371%2Fjournal.pone.0031883.
28. Lal A, Ikeda T, French N, Baker MG, Hales S. Climate Variability, Weather and Enteric Disease Incidence in New Zealand: Time Series Analysis. Plos One. 2013;8(12):e83484.
29. Hoque M, Hope V, Scragg R, Kjellström T. Children at risk of giardiasis in Auckland: a case–control analysis. Epidemiol Infect. 2003;131(01):655-62.
30. Hoque M, Hope V, Scragg R, Baker M, Shrestha R. A descriptive epidemiology of giardiasis in New Zealand and gaps in surveillance data. N Z Med J. 2004;117(1205):U1149.
31. Sagebiel D, Weitzel T, Stark K, Leitmeyer K. Giardiasis in kindergartens: prevalence study in Berlin, Germany, 2006. Parasitol Res. 2009;105(3):681-7.

ACCEPTED MANUSCRIPT

32. Stefanogiannis N, McLean M, Van Mil H. Outbreak of cryptosporidiosis linked with a farm event. N Z Med J. 2001;114(1144):519-21.
33. Grinberg A, Pomroy WE, Squires RA, Scuffham A, Pita A, Kwan E. Retrospective cohort study of an outbreak of cryptosporidiosis caused by a rare *Cryptosporidium* parvum subgenotype. Epidemiol Infect. 2011;139(Special Issue 10):1542-50.
34. Ng J, Eastwood K, Durrheim D, Massey P, Walker B, Armson A, et al. Evidence supporting zoonotic transmission of *Cryptosporidium* in rural New South Wales. Exp Parasitol. 2008;119(1):192-5.
35. Chalmers RM, Smith R, Elwin K, Clifton-Hadley FA, Giles M. Epidemiology of anthroponotic and zoonotic human cryptosporidiosis in England and Wales, 2004-2006. Epidemiol Infect. 2011;139(5):700--12.
36. Li X, Atwill ER, Dunbar LA, Tate KW. Effect of daily temperature fluctuation during the cool season on the infectivity of *Cryptosporidium* parvum. Appl Environ Microbiol. 2010;76(4):989-93.
37. Lake IR, Bentham G, Kovats RS, Nichols GL. Effects of weather and river flow on cryptosporidiosis. J Water Health. 2005 Dec;3 (4):469-74.
38. Graczyk TK, Knight R, Tamang L. Mechanical transmission of human protozoan parasites by insects. Clinical Microbiology Reviews. 2005;18(1):128-32.
39. Wheeler J, Sethi D, Cowden J, Wall P, Rodrigues L, Tompkins D, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. BMJ. 1999;318:1046 - 50.
40. Lal A, Baker MG, French NP, Dufour M, Hales S. The epidemiology of human salmonellosis in New Zealand, 1997-2008. Epidemiol Infect. 2011:1-10.

Table 1. Retrospective multivariate modelling for giardiasis for each independent variable (Z_i), its associated posterior coefficient estimate ($\hat{\beta}_k$), the expected change in the rate of notified case ($e^{\hat{\beta}_k}$) with the corresponding 95% credible interval. The reference categories are the percentage of the population aged between 4 and 65 years of age, the percentage of the population with European ethnicity (including New Zealand European), the most affluent CAUs, good drinking water quality and rural CAUs

Explanatory Variable (Z_i)	Model coefficient ($\hat{\beta}_k$)	Incidence Risk Ratio ($e^{\hat{\beta}_k}$) (95% CI)
Average Rainfall (mm)	0.035	1.04 (0.98, 1.09)
Average Temperature (°C)	-0.023	0.98 (0.93, 1.03)
Percent aged 4-65 Years	Reference	1.00
Percent aged < 4 Years	0.041	1.04 (1.02, 1.06)
Percent aged ≥ 65 Years	-0.0083	0.99 (0.98, 1.00)
Dairy Cattle Density (cows/ km^2)‡	-0.096	0.91 (0.85, 0.98)
Deer Density‡	-0.019	0.98 (0.89, 1.08)
Poultry Density‡	-0.048	0.95 (0.91, 1.00)
Pig Density‡	0.034	1.04 (0.93, 1.15)
Beef Cattle Density‡	0.011	1.01 (0.94, 1.09)
Percent European	Reference	1.00

ACCEPTED MANUSCRIPT

Percent Māori	-0.013	0.99 (0.98, 0.99)
Percent Pacific Islander	-0.025	0.98 (0.97, 0.98)
Percent Asian	-0.021	0.98 (0.97, 0.99)
Percent Middle Eastern/Latin American/African	0.012	1.01 (0.99, 1.04)
Good Drinking Water Quality	Reference	1.00
Intermediate Drinking Water Quality	-0.029	0.97 (0.91, 1.04)
Poor Drinking Water Quality	-0.08	0.92 (0.85, 1.00)
Unknown Drinking Water Quality	-0.019	0.98 (0.88, 1.09)
Rural residence	Reference	1.00
Urban residence	-0.077	0.93 (0.78, 1.09)
Population Density‡	-0.17	0.84 (0.77, 0.92)
Area Socio-Economic Deprivation (Least deprived)	Reference	1.00
Area Socio-Economic Deprivation (Intermediate)	-0.063	0.94 (0.87, 1.01)
Area Socio-Economic Deprivation (Most deprived)	-0.00067	1.00 (0.88, 1.14)

‡ Density is defined as the number of animals or people per square kilometre. For density variables the IRR represents the effect of a one $\log_{10}$ change

Table 2. Retrospective multivariate modelling for cryptosporidiosis for each independent variable (Z_i), its associated posterior coefficient estimate ($\hat{\beta}_k$), the expected change in the rate of notified case ($e^{\hat{\beta}_k}$) with the corresponding 95% credible interval. The reference categories are the percentage of the population aged between 4 and 65 years of age, the percentage of the population with European ethnicity (including New Zealand European), the most affluent CAUs, good drinking water quality and rural CAUs

Explanatory Variable (Z_i)	Model coefficient ()	Incidence Risk Ratio ($\hat{\beta}_k$) (95% CI)
Average Rainfall (per increase in 1 mm)	0.035	1.04 (0.97, 1.11)
Average Temperature (per increase in 4°C)	-0.076	0.93 (0.87, 0.98)
Percent aged 4-65 Years	Reference	1.00

ACCEPTED MANUSCRIPT

Percent aged < 4 Years	0.036	1.04 (1.02, 1.06)
Percent aged ≥ 65 Years	-0.012	0.99 (0.98, 1.00)
Dairy Cattle Density‡ (cows/km^2)	0.12	1.13 (1.05, 1.22)
Deer Density‡	-0.071	0.93 (0.85, 1.03)
Poultry Density‡	0.023	1.02 (0.98, 1.07)
Pig Density‡	-0.047	0.95 (0.85, 1.06)
Beef Cattle Density‡	-0.021	0.98 (0.90, 1.07)
Percent European	Reference	1.00
Percent Māori	-0.01	0.99 (0.98, 1.00)
Percent Pacific Islander	-0.019	0.98 (0.97, 0.99)
Percent Asian	-0.021	0.98 (0.97, 0.99)
Percent Middle Eastern/Latin American/African	-0.0046	1.00 (0.97, 1.02)
Good Drinking Water Quality	Reference	1.00
Intermediate Drinking Water Quality	-0.023	0.98 (0.90, 1.06)
Poor Drinking Water Quality	-0.032	0.97 (0.88, 1.07)
Unknown Drinking Water Quality	-0.13	0.88 (0.77, 1.00)
Rural residence	Reference	1.00
Urban residence	-0.15	0.86 (0.74, 1.00)
Population Density‡	-0.30	0.74 (0.69, 0.80)
Area Socio-Economic Deprivation (Least deprived)	Reference	1.00
Socio-Economic Deprivation (Intermediate)	-0.039	0.96 (0.88, 1.05)
Socio-Economic Deprivation (Most Deprived)	-0.065	0.94 (0.80, 1.10)

‡Density is defined as the number of animals or people per square kilometre. For density variables the IRR represents the effect of a one $\log_{10}$ change

ACCEPTED MANUSCRIPT

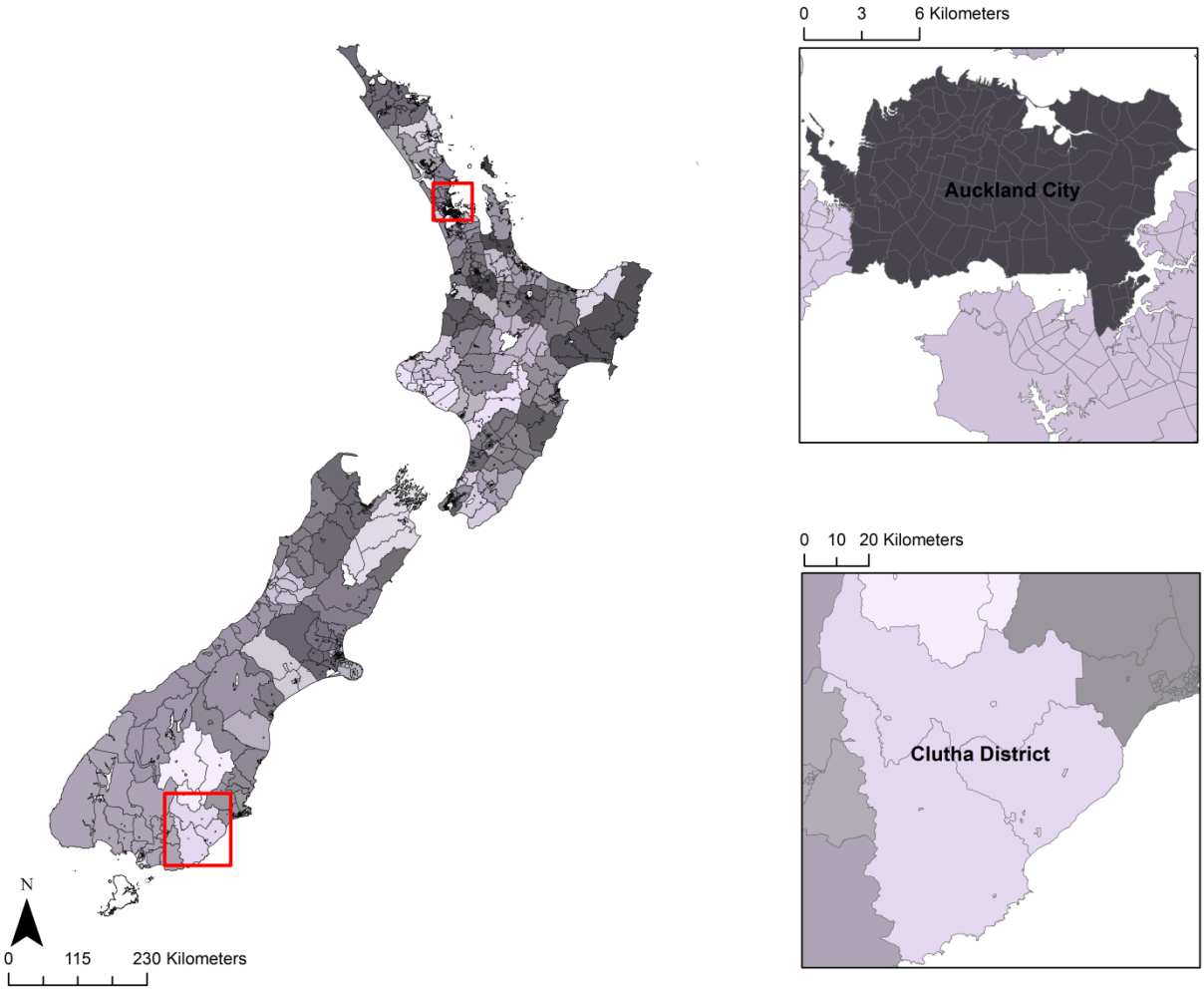


Fig. 1. Distribution of 2006 Census Area Units (CAU) grouped into Territorial Authorities (TA) (identified using different shades of grey) used for outbreak detection regions. Inset maps show Regions selected for detailed descriptive analysis of patterns: Auckland City (urban) and Clutha District (rural).

ACCEPTED M

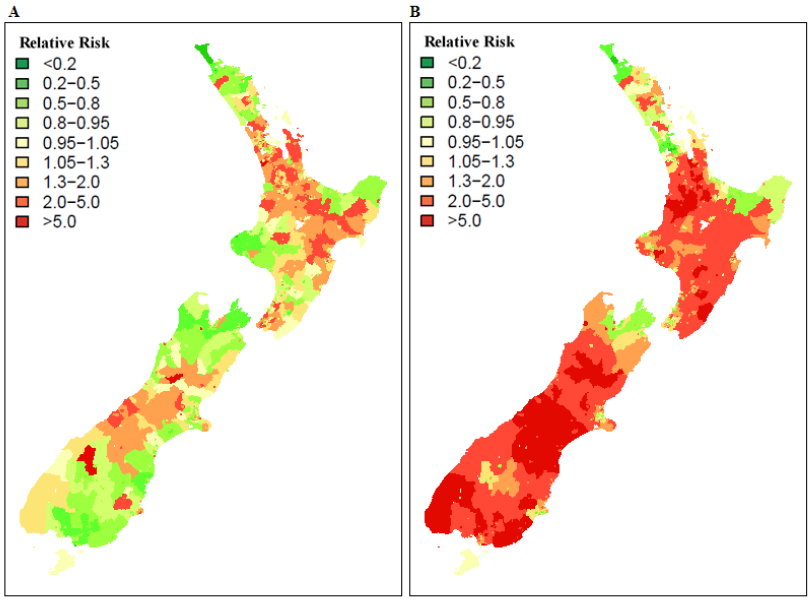


Fig. 2. Spatial Distribution of relative risk for (A) giardiasis (B) cryptosporidiosis across the whole country. The relative risk is interpreted as the risk of disease in each CAU as compared to the average risk across all CAUs. A value greater than 1.0 implies a higher than average risk, whereas a value less than 1.0 implies a lower than average risk.

ACCEPTED MANU

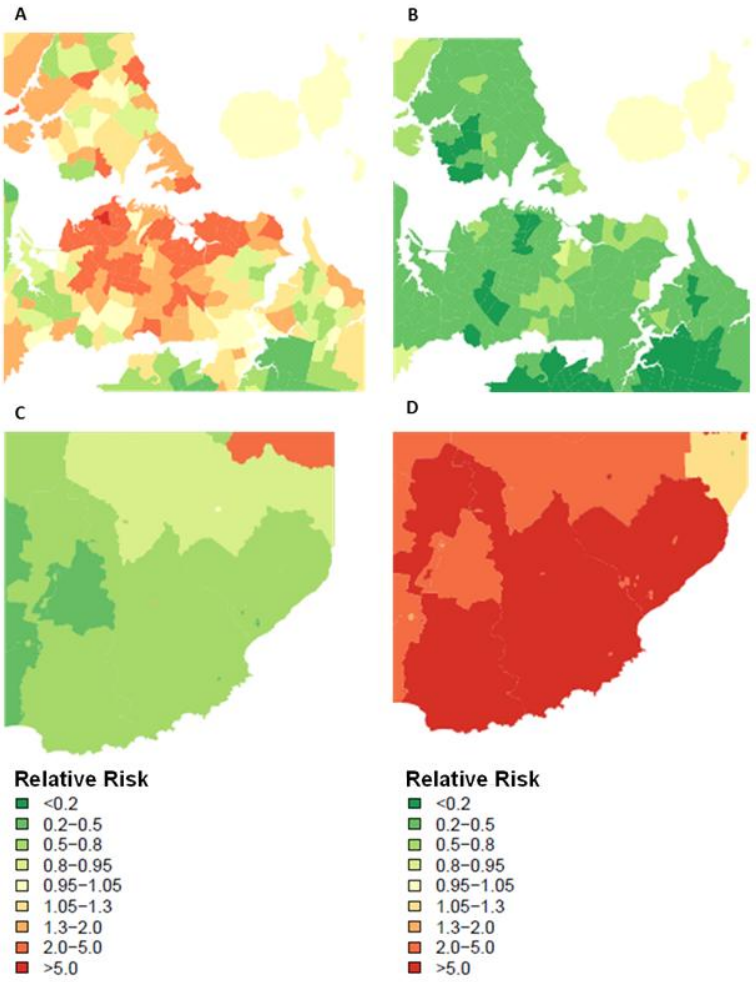


Fig. 3. Spatial distribution of (A) relative risk for giardiasis in Auckland City (B) relative risk for cryptosporidiosis in Auckland City (C) relative risk for giardiasis in Clutha District (D) relative risk for cryptosporidiosis in Clutha District.

ACCEPTED MA

ACCEPTED MANUSCRIPT

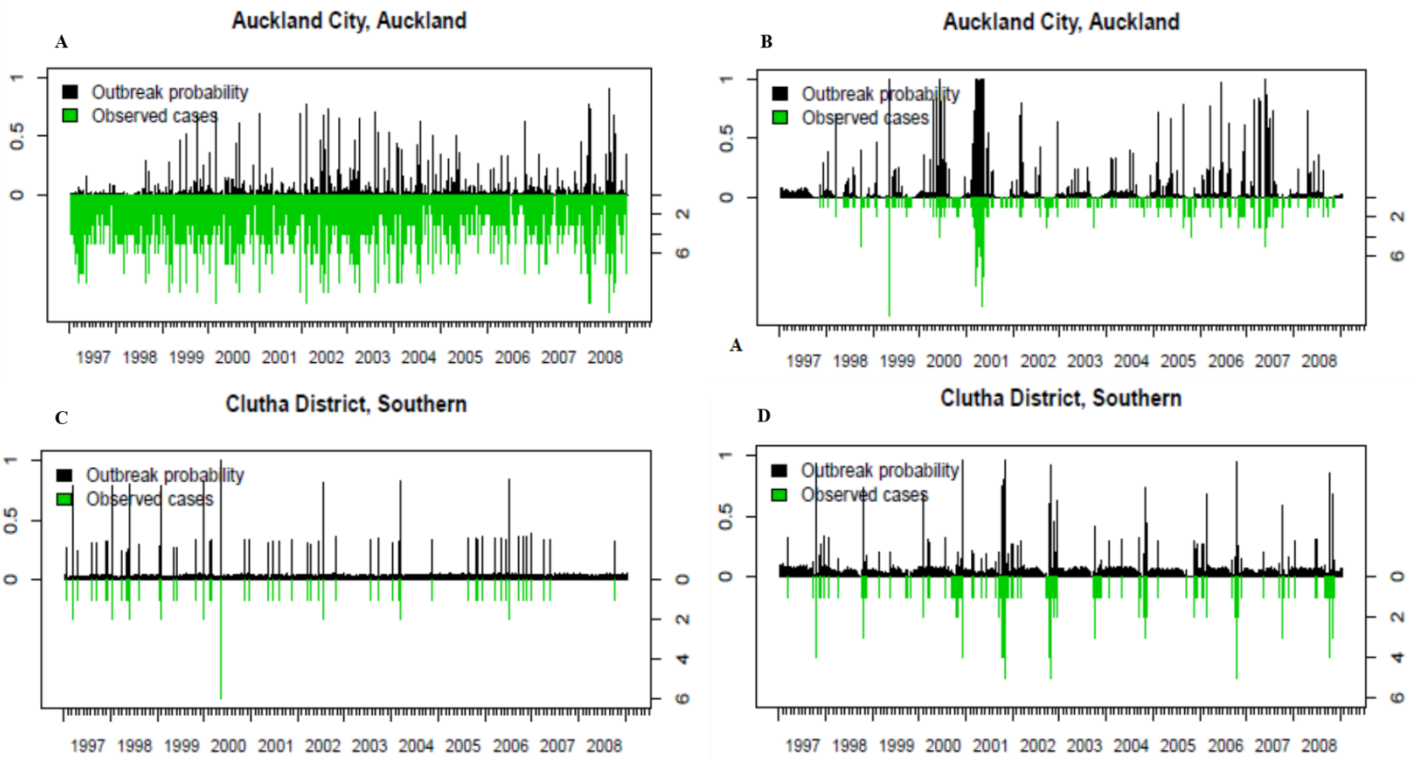


Fig. 4. The posterior probability of a (A) localised giardiasis outbreak in Auckland City (B) localised cryptosporidiosis outbreak in Auckland City (C) localised giardiasis outbreak in Clutha District (D) localised cryptosporidiosis outbreak in Clutha District. The y axis represents the probability (black) and the number of cases (green) and the x axis is month.

ACCEPTED MANUSCRIPT