An island at risk: geographic epidemiologic analysis of the 1848 – 1850 cholera epidemic in Ireland

April VP Clyburne-Sherin[1,2], Olaf Berke[1,2], Eyal Klement[3], and Jan Sargeant[1,2]

[1]Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, Ontario, CANADA, N1G 2W1
[2]Centre for Public Health and Zoonoses, University of Guelph, Guelph, Ontario, CANADA, N1G 2W1
[3]Koret School of Veterinary Medicine, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, ISRAEL, 76100

ABSTRACT

Background
Cholera is considered a re-emerging disease whose spread remains difficult to predict. One way to improve prediction of cholera epidemics is to learn from the dynamics of past epidemics. This study investigates the 1848 – 1850 cholera epidemic in Ireland during the Irish Famine. The aim of this study is to analyse and visualize the distribution of cholera during this epidemic and the role of demographics and famine in cholera attack rates and case fatality rates.

Methods
Historical cholera data were retrieved from a government report by Ireland's Commissioners of Health in 1852, Ireland's 1851 census, and a report by the Central Relief Committee in 1852. These data were used to create maps to describe the spatial and temporal patterns of cholera, demographics, and famine relief. Higher than expected attack rates or case fatalities were detected using the spatial scan test. Associations between (1) cholera attack rates and (2) case fatality rates with demographic and famine data were modelled using spatial generalized linear mixed models.

Results
Cholera cases were first reported on the North-Eastern coast of Ireland, with some sporadic spread to follow on the Central-Eastern coast, and the possibility of a separate arrival of cholera from the South-Western coast. People living in Galway county had 3.42 (95% CI 2.61 – 4.49) times greater risk of contracting cholera than those in other counties, and Antrim, King's, Waterford, Kilkenny, Tipperary, Wexford, Cork, and Carlow counties had 1.14 (95% CI 1.11 – 1.17) to 1.65 (95% CI 1.48 – 1.84) times greater risk. People who had contracted cholera were 1.15 (95% CI 1.10 – 1.20) times more at risk of dying from their infection in King's, Queen's, Westmeath, Kildare, Longford, Tipperary, Roscommon, Kilkenny, Carlow, and Meath counties, and 1.13 (95% CI 1.04 – 1.22) to 1.31 (95% CI 1.15 – 1.48) times more at risk in Leitrim, Sligo, and Waterford counties. Of the 32 counties of Ireland, 18 were identified as part of clusters of either elevated attack rates or case fatality rates. However, these clusters overlapped only by 5 counties and covered otherwise different regions of the study area. Cholera attack rates were not explained by demographics and famine but by factors not included in this study, whereas case fatality rates were significantly associated with families per house.

Conclusions
This outbreak provides insight into the natural history of cholera before there were effective public health interventions or public understanding of transmission. Although everyone was at risk of contracting cholera during this outbreak, some were more likely to die of their infection. This study provides evidence that people that were housed more densely had better chance of surviving their infection. This seemingly counterintuitive result is suggested here to be an indicator of importance of access to health infrastructure.

Keywords: Spatial epidemiology, Statistics, Cholera, Ireland

INTRODUCTION

Cholera is one of the foremost water-borne diseases of citizens in developing countries without adequate access to clean water or sanitation facilities (Jutla et al., 2013; Fac, 2014). Every year, 3 to 5 million people worldwide are infected by cholera, of which more than 100,000 die (Fac, 2014). Cholera is a re-emerging disease whose spread remains difficult to predict (Jutla et al., 2013).

One way to improve our ability to predict cholera epidemics is to learn from the dynamics of past epidemics. Studying historical cholera epidemics improves our understanding of the natural history of cholera before there were effective public health interventions or public understanding of transmission (Chan et al., 2013). Recent examples of studies that revisited cholera epidemics include Bingham and colleagues (Bingham et al., 2004) who studied the role of the water supply in the London outbreak of 1849, and Kuo and colleagues (Kuo and Fukui, 2007) who identified the changing geographical patterns of cholera in Fukushima from 1882 and 1895. Here, the 1848 – 1850 cholera epidemic in Ireland is investigated. This epidemic was part of the second cholera pandemic, during which more than 45,698 cases and 19,325 deaths were reported to Ireland's Central Board of Health (Rep, 1852). It began during the Irish famine (1845 – 1852), when, between the censuses of 1841 to 1851, the population of Ireland dropped by 1,622,739 people through starvation, disease, and emigration, to a total population of 6,552,385 (Cen, 1856) (7) (Figure 1).

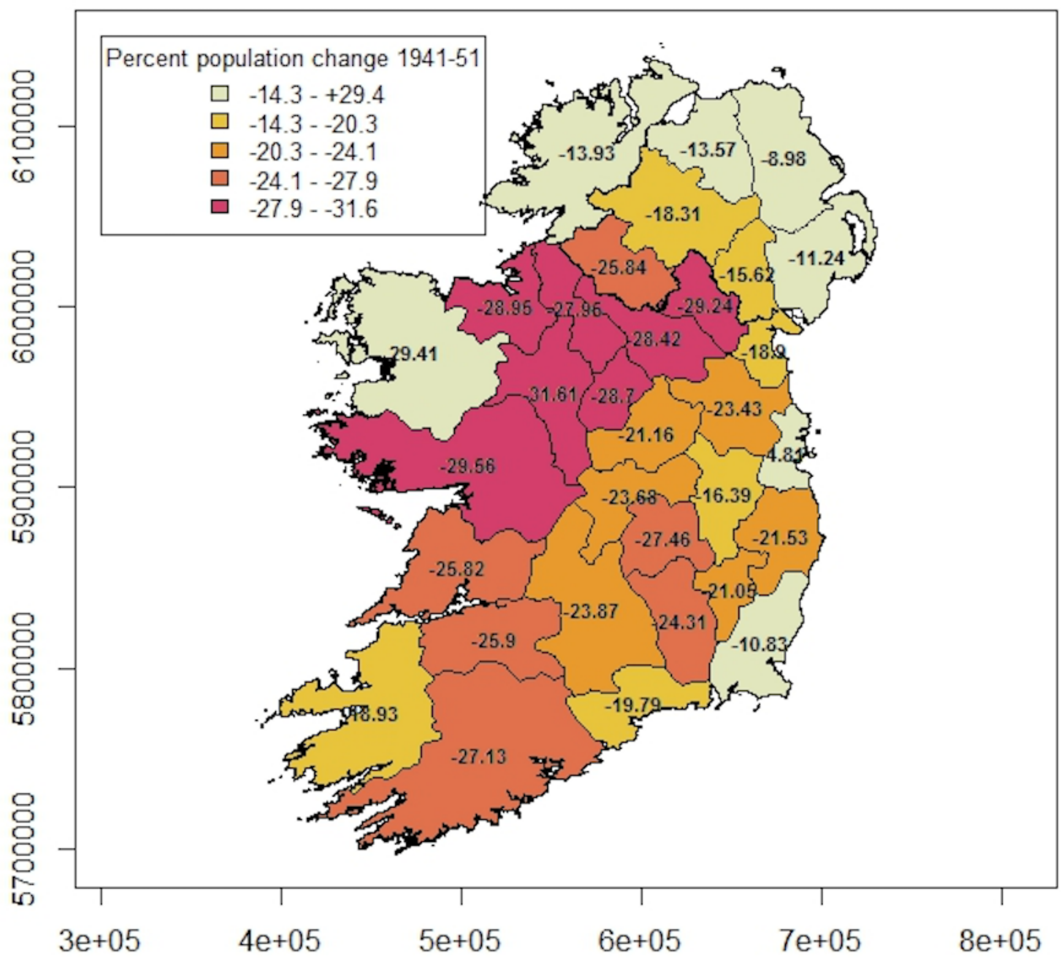

Figure 1. Percent population change by county in Ireland from 1841 – 1851.

Epidemiology of the cholera epidemic of 1848 – 1850 in Ireland

The second cholera pandemic began in 1829 in India, and first arrived in Ireland in Dublin in March, 1832. This first arrival of cholera in Ireland spread to the principal towns, but subsided quickly. The famine of Ireland began with the potato blight in 1845, which led to migration of Irish people to cities and workhouses in search of work and food. The first reported case of this epidemic was a man just arrived from Edinburgh and took place on December 4, 1848 in a workhouse in Belfast.

Demographics and famine during the epidemic

Overcrowding (Kuo and Fukui, 2007; Tauxe et al., 1988), poverty (Fac, 2014), migration (Arnold, 1993), inadequate health infrastructure (Rep, 1852), famine (Tauxe et al., 1988; Arnold, 1993) and malnutrition (Arnold, 1993; Young and Jaspars, 1995) can increase the risk of cholera, and were all issues in Ireland during this cholera epidemic (Rep, 1852). In 1852, Ireland's Commissioners of Health presented a report containing data and commentary on cholera and other epidemics between 1846 – 1850, titled *Report of the Commissioners of Health, Ireland, on the epidemics of 1846 to 1850: presented to both Houses of Parliament by Command of her Majesty* (Rep, 1852).

Rural living and health infrastructure

Within this report, the Commissioners described the state of Ireland's health infrastructure as "destitute of sufficient funds, and dispensaries" and unfit to provide adequate medical aid, especially to rural inhabitants "scattered over districts" (Rep, 1852). Starting in February of 1847, a spike in applications to the Board of Health to improve the inadequate health infrastructure through the establishment of temporary hospital accommodation was noted by the Commissioners. However, of 576 applications for hospital accommodation relief, 203 were refused (Rep, 1852). The Commissioners themselves commented on the higher than average case fatality rates during the study period, and had requested explanations for the high fatalities from all Medical Officers at locations where it exceeded 15%. The Medical Officers responded, referring to inadequate health infrastructure as a dominant reason for high case fatality rates at their hospitals, and repeatedly citing overcrowded hospitals, long distances between hospitals, or poor construction of hospitals as explanations. Without the temporary hospital accommodations, patients were arriving weakened "from the long distance many had to travel to hospital", and the Medical Officers were able to admit "only the worst cases" (Rep, 1852). Therefore, there was reason to suspect that inadequate health infrastructure and hospital overcrowding may have played a role in this cholera epidemic, as it has in others (Kuo and Fukui, 2007; Arnold, 1993).

Famine and education

This cholera epidemic may have been influenced by the Irish famine in multiple ways, both directly and indirectly. Famine can have an indirect impact on disease through a breakdown of social relationships, vagrancy, and overcrowding in unsanitary housing (Arnold, 1993). Famine can also directly cause malnutrition or the consumption of foods not normally consumed by the population, known as "famine foods", which were sometimes improperly prepared or nutritionally inadequate, despite efforts to educate the public on their preparation (Arnold, 1993). Examples of famine foods in place of the potato were described by the Commissioner's report by a Medical Officer who complained that disease in Ireland was "caused and kept up by bad food... the consumption of meal or rice, but more particularly Indian meal, in a raw or badly cooked state" (Rep, 1852). Medical Officers in the Commissioner's report were almost unanimous in citing the famine as the main reason for high case fatality rates during this period in the section "Observations of Medical Officers relative to the rate of Mortality". Their patients were generally "worn down by cold and famine", and their deaths were attributed to the "starvation" or "want of proper food" prior to infection (Rep, 1852).

Poverty and overcrowding

Medical Officers in the Commissioner's report also attributed their patients' deaths to the "destitution" and the "debilitated condition" of the poor prior to contracting cholera (Rep, 1852). John Snow observed that cholera was more fatal "amongst the vagrant class, who lived in [a] crowded state" in London in 1832 (Snow, 1855). As in previous cholera epidemics (Fac, 2014; Kuo and Fukui, 2007; Tauxe et al., 1988), poverty and overcrowded housing may have played a role.

Study goal and objectives

Cholera is considered a re-emerging disease whose spread remains difficult to predict. One way to improve prediction of cholera epidemics is to learn from the dynamics of past epidemics by studying the spatio-temporal pattern of an epidemic and identifying possible predictors for the pattern such as overcrowding (Kuo and Fukui, 2007; Tauxe et al., 1988), poverty (Fac, 2014), inadequate health infrastructure (Rep, 1852), famine (Tauxe et al., 1988; Arnold, 1993) and malnutrition (Arnold, 1993; Young and Jaspars, 1995). The goal of this study was to gain insight in the spatio-temporal spread of cholera by applying modern geographical epidemiologic methods to this historic cholera epidemic. Specific objectives were: (1) to describe the spatio-temporal pattern of the attack rates and case fatality rates at town, county and kingdom levels; (2) to detect and locate clusters of elevated attack rates and case fatality rates; and (3) to identify and estimate the association of attack rates and case fatality rates with demographic and famine related predictors.

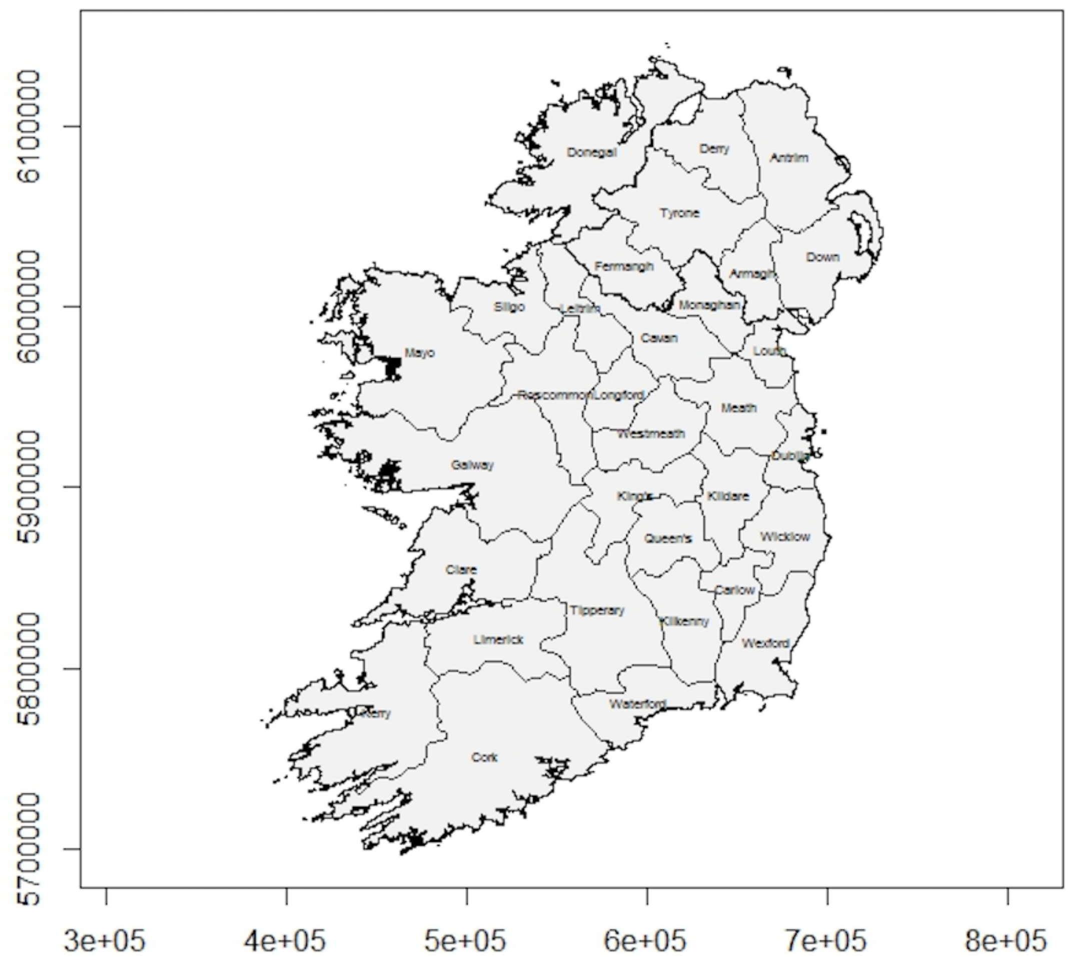

Figure 2. County boundaries of Ireland (1848 – 1850).

METHODS

Study area

During the study period of 1848 – 1850, The Kingdom of Ireland, hereafter known as Ireland, was subdivided into 4 provinces, 32 counties, and 320 baronies (Cen, 1856) and covered an area of 84,421

km^2 (Nolan, 2014) (Figure 2).

Data collection

Historical data regarding the cholera outbreak in Ireland from 1848 – 1850 was recorded in a government report by Ireland's Commissioners of Health in 1852, titled *Report of the Commissioners of Health, Ireland, on the epidemics of 1846 to 1850: presented to both Houses of Parliament by Command of her Majesty* (Rep, 1852). The report included data from 136 towns of Ireland, i.e. all towns with a population greater than 2000 people (Appendix 1). The report included the monthly total number of cholera cases reported to the Central Board of health for all of Ireland (Appendix 2) and monthly total number of cholera cases and deaths for six cities: Belfast, Cork, Dublin, Galway, Limerick, and Waterford (Appendix 3).

Data level	Historical cholera and demographic data	n	Source
Town level data	Population	136	6
	Total number of cholera cases	136	6
	Total number of cholera deaths	136	6
	Rate of mortality	136	6
	Date of first cholera case	136	6
	Date of last cholera case	136	6
	Percent persons over age 5 who cannot read or write	136	7
City level data	Total monthly cholera cases	78	6
	Total monthly cholera deaths	78	6
	Rate of mortality	78	6
County level data	Percent persons over age 5 who cannot read or write	32	7
	Percent persons recorded as having health conditions	32	7
	Average number of families per 100 houses	32	7
	Number of persons per square statute mile	32	7
	Percent of families living in 4th class housing ("mud huts")	32	7
	Percent change in population from the 1941 census to the 1951 census	32	7
	Quaker famine relief aid in total tons of food from 1846 - 1848	32	13
Kingdom level data	Total monthly cholera cases	21	6

Table 1. Historical cholera and demographic data collected from the Report of the Commissioners of Health, Ireland, on the epidemics of 1846 to 1850, the Census of Ireland for the Year 1851, and Extracts from Transactions of the Central Relief Committee of the Society of Friends during the Famine in Ireland in 1846 and 1847.

The *Census of Ireland for the Year 1851* contains data regarding the demographics of Ireland near the time of the cholera outbreak (Cen, 1856). The historical report *Extracts from Transactions of the Central Relief Committee of the Society of Friends during the Famine in Ireland in 1846 and 1847* contained additional data regarding the degree of famine relief provided to each county (Ext, 1852). The historical data collected is presented in Table 1. The borders of Ireland and Ireland's counties were created by merging the shapefiles (GADM database version 2.0, available from http://www.gadm.org) of the current regions of Northern Ireland and Ireland using QGIS (version 2.6.0, Open Source Geospatial Foundation Project, available from http://qgis.osgeo.org). The merged map of county boundaries and town locations were compared with the map of Ireland from the *Report of the Commissioners of Health, Ireland, on the epidemics of 1846 to 1850: presented to both Houses of Parliament by Command of her Majesty* published in 1852 to ensure the study boundaries used were historically correct.

Analyses

For spatial analysis, the data as recorded in the Report of the Commissioners of Health were used to calculate the following epidemiological measures by town, county, and kingdom:

Attack rate = cholera cases / population at risk

Mortality rate = cholera deaths / population at risk

Case fatality rate = cholera deaths / cholera cases

Epidemic day = the number of days since the day of the first reported cholera case of the epidemic (epidemic day 1 = December 4, 1848)

The population at risk was defined as the total population of the town, county, or kingdom for which the measures were calculated.

Analysis of the cholera epidemic at the town level

Spatial

Attack rates, mortality rates, and case fatality rates were calculated for each town. These measures were then used to visualize the spatial patterns of the epidemic at the town level. Cholera occurrence at the town level was visualized using a dot map, comparing the distribution of towns that reported cholera cases and with those that did not. The total number of cholera cases, the total number of cholera deaths, attack rates, and case fatality rates per town were compared using proportional symbol maps.

Temporal

The epidemic day of the first reported cholera case and the epidemic day of the last reported cholera case for each town were extracted. The spatial pattern of the cholera epidemic spread in time was visualized using an isopleth risk map of the epidemic day of the first reported cholera case per town using ordinary, model-based kriging. Kriging interpolates continuous data based on the weighted average of neighbouring observations (Berke, 2004). Spatial dependence structures were modelled to best fit the data using semivariograms. An exponential semivariogram model with nugget effect was fit to the data using the robust estimator (Cressie and Hawkins, 1980).

Analysis of the cholera epidemic at the county and kingdom level

Spatial

Risk maps

Regional data were interpolated into a continuous surface to create isopleth maps. Isopleth risk maps of Empirical Bayesian smoothed cholera attack rates, mortality rates, and case fatality rates were created using ordinary, model-based kriging (Berke, 2004) (14). Spatial dependence structures were modelled to best fit the data using exponential semivariograms with a nugget effect. Estimator type for the semivariogram was selected as robust because for the spatial sample size was low with 32 counties.

Cluster detection

The spatial scan statistic was applied to detect "clusters" in cholera attack rates and case fatality rates using SaTScan software (Hjalmars et al., 1996; Kulldorff, 1997). A cluster was defined as an area where the attack rate or case fatality rate of the population was significantly higher than for the population outside the area where the cluster resides. The spatial scan test used circular scanning windows to visit Ireland's county centroids, increasing the radius of the scanning window and performing a likelihood ratio test to see if the attack or case fatality rates inside the scanning window were significantly greater than expected based on rates outside of the scanning window. A limit of 30% of the total population per scanning window was selected to limit the area of potential clusters. A Poisson distribution was used for the spatial scan test, with 999 Monte Carlo simulations to estimate p-values ($\alpha = 0.05$).

Temporal

An epidemic curve of the total monthly cases in Ireland from December 1848 to August 1850 was plotted to assess the temporal distribution of the epidemic at the kingdom level.

Variable	Min.	Lower quartile	Mean	Median	Upper quartile	Max.
Percent illiterate	20.0	38.0	47.0	46.5	57.5	74.0
Percent with health conditions	0.9	1.2	1.6	1.4	1.9	2.6
No. families per 100 houses	105	107	110	109	113	121
Percent families in 4th class houses	3.5	9.7	13.4	12.0	16.8	31.6
Famine relief in food (tons)	8.0	38.7	244.7	95.5	300.7	1866.0
No. persons per km^2	47.1	60.1	77.2	70.7	79.6	162.9
Attack rate	0	0.014	0.024	0.026	0.031	0.086
Case fatality rate	0.27	0.39	0.45	0.45	0.50	0.60

Table 2. Descriptive statistics of county level dependent and explanatory variables.

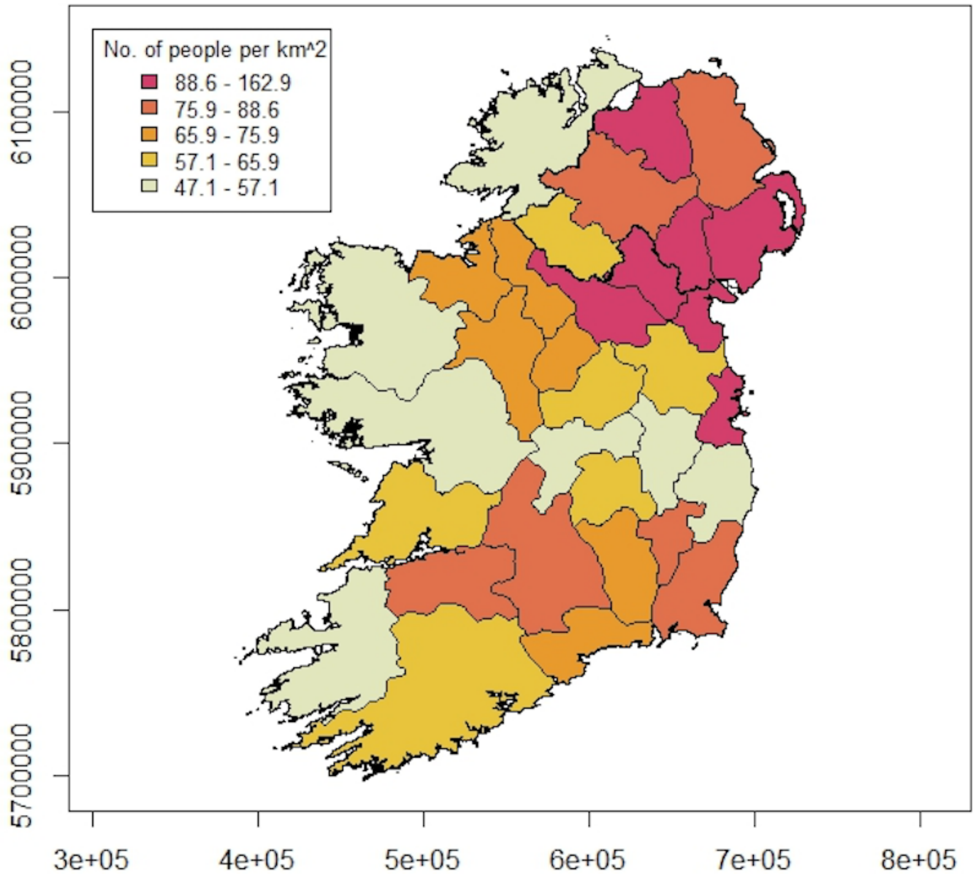

Figure 3. Average number of people per square kilometre by county in Ireland (1851).

Descriptive exploration of Ireland's demographics at the county level

The distribution of Ireland's demographic and famine covariates were explored using descriptive statistics (Table 2). The spatial variation in demographics and famine relief was visualized using the choropleth maps below.

Rural living and population density

Figure 3 illustrates the average number of people per square kilometre as a measure for population density and urbanization for each county. The North-Eastern coast of Ireland had high population density, with the rest of Ireland having a lower population density.

Population Health

Figure 4 illustrates the percent of the population reported as "sick" in the 1851 census, which included those who were "Deaf and Dumb", "Blind", "Lunatic", "Idiotic", "Lame and Decrepit", and "Sick from other causes" as a proxy measure for population health for each county. The South-Western region of Ireland had the highest percentage of people with health conditions.

Education

Figure 5 illustrates the percent population over the age of 5 "who cannot read nor write" as reported in the 1851 census as a proxy measure for education for each county. The Western coast of Ireland had the

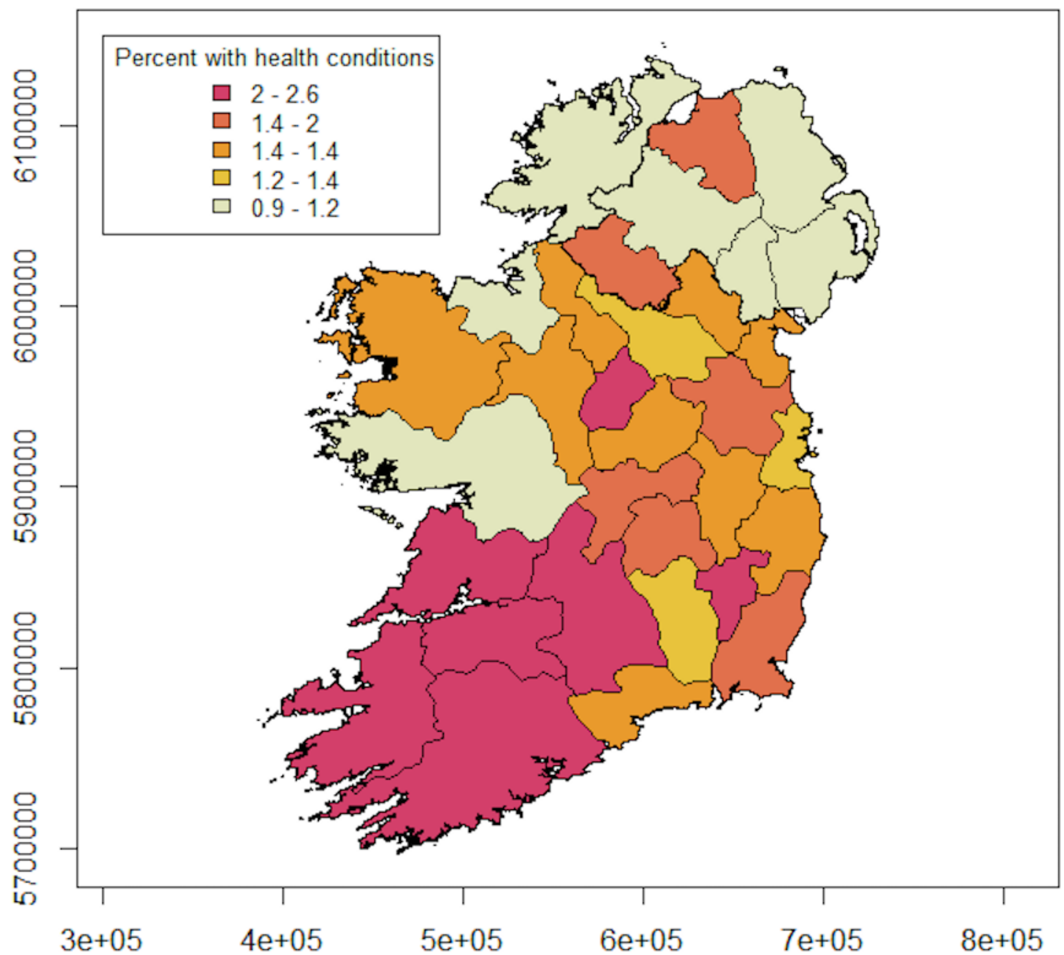

Figure 4. Percent of the population with health conditions by county in Ireland (1851).

highest percentage of illiterate people, with Central and Eastern coast regions were inhabited by a more literate population.

Poverty

Figure 6 illustrates the percent of families living in 4th class housing, consisting of ”mud huts”, as a proxy measure for poverty for each county. Both the South-Western coast and the Central-Western coast of Ireland had high percentages of families living in ”mud huts”, whereas people in the North-Eastern region had a lower percentage.

Overcrowding

Figure 7 illustrates the average number of families per 100 houses as a proxy measure for overcrowding for each county. The South-Western coast of Ireland had high numbers of families per 100 houses, with the South generally had higher numbers of families per 100 houses than the North of Ireland.

Famine

Figure 8 illustrates the total Quaker famine relief measured by tons of food received as a proxy measure for the impact of famine for each county. The Western coast of Ireland received the most famine relief, with the Central and Eastern regions having received much less relief.

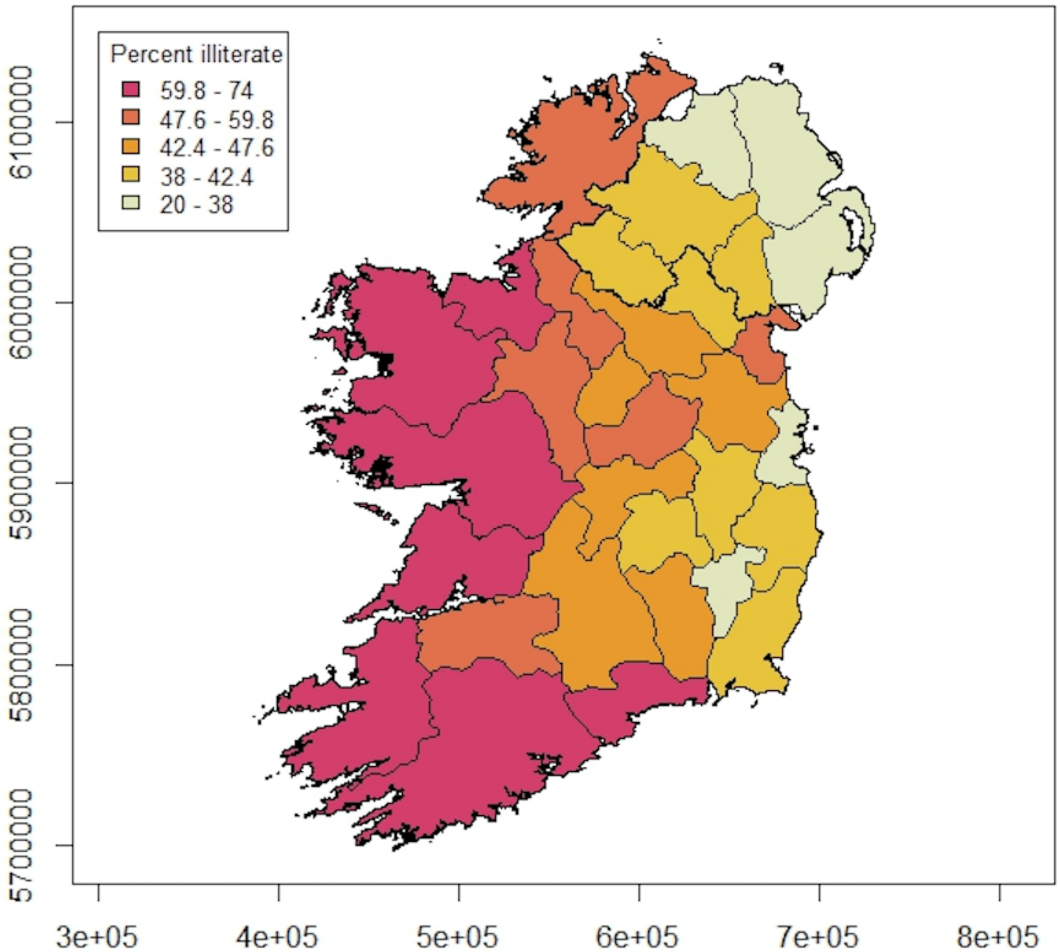

Figure 5. Percent population over age 5 who cannot read or write by county in Ireland (1851).

Modelling attack rates and case fatality rates

Two models were fit to estimate the association of the cholera epidemic with demographic and famine data. Both the cholera attack rates and case fatality rates were separately modelled using spatial Poisson regression models with random effects to account for spatial dependence in the data (Dormann et al., 2007). A spatial Poisson model was also fit to assess whether attack rates were associated with case fatality rates. Spherical dependence structures were used without a nugget effect, and then fit by penalized quasi-likelihood estimation using the R function glmmPQL (Breslow and Clayton, 1993). These models estimated the associations between (1) attack rates per county or (2) case fatality rates per county with the county-level demographic and famine covariates: (i) percent illiterate people as an indicator for education; (ii) percent of people with health conditions as an indicator for population health; (iii) average number of families per 100 houses as an indicator for overcrowding; (iv) number of persons per square kilometre as an indicator for population density; (v) percent families in 4th class houses as an indicator for poverty; and (iv) famine relief in tons of food. Spearman's rank correlation coefficients were estimated between all county level dependent and explanatory variables to assess collinearity, with a critical value of $|r| > 0.7$ (Dormann et al., 2013). Models were fit using backward selection: starting with the full set of explanatory variables in the models, and removing those variables which did not contribute significantly to the models ($\alpha = 0.20$). When remaining model parameter(s) changed more than 10% upon removal

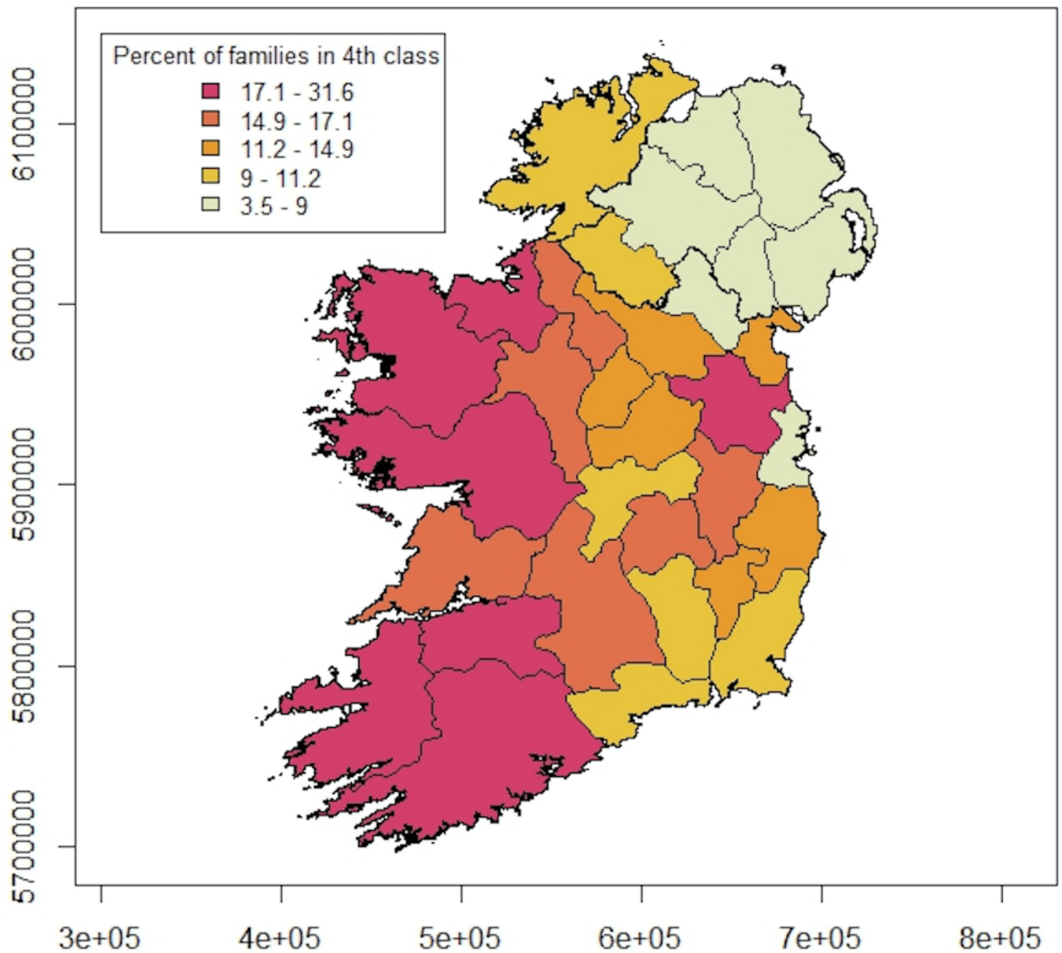

Figure 6. Percent of families living in 4th class housing by county in Ireland (1851).

they were judged as confounders and were retained (Dohoo et al., 2003). Two-way interaction terms were added to the resulting model, and included if significant (α = 0.20). Normal quantile-quantile plots of deviance residuals and observed versus predicted plots were used to evaluate the fit of the models. All analyses were conducted in R (version 3.1.3, Foundation for Statistical Computing, available from http://www.R-project.org).

RESULTS

Analysis of the cholera epidemic at the town level

Spatial

Cholera affected 103 of the 136 major towns that were included in the Commissioner's report (Figure 9). Of those towns included in the report, the average number of total cholera cases per town was 245 cases (range 0 – 3813; median 90; IQR 2 – 245.8) and the average number of total cholera deaths per town was 109 deaths (range 0 – 1664; median 36; IQR 1 – 114.2). The total number of reported cases and reported deaths per town shows that the most populous cities, Dublin, Cork, and Belfast, had both the greatest number of total cases and the greatest number of total deaths (Figures 10 and 11). The average attack rate per 100,000 people was 2896 cases or 2.9% per town's population (range 0 – 34.16%; median 1.69%; IQR 0.05 – 4.17%) (Figure 12). The average mortality per 100,000 people was 1290 cholera deaths or

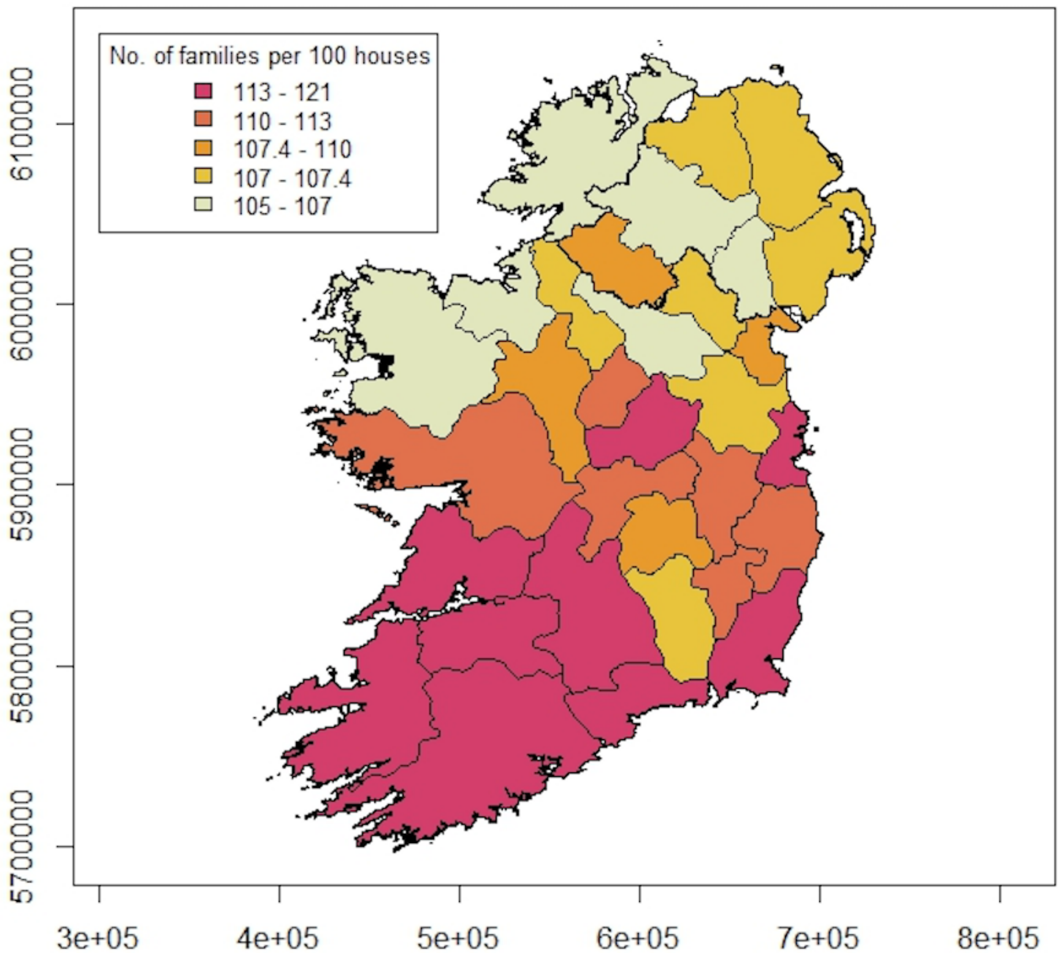

Figure 7. Average number of families per 100 houses by county in Ireland (1851).

1.3% per town's population. The average case fatality rate for the towns in the report was 47% (range 0 – 100%; median 47.17%; IQR 39.05 – 54.12%) (Figure 13). Gort, Ballinasloe, and Kinsale had both the highest attack rates and highest mortality rates.

Temporal

The average epidemic day of first reported cholera case of infected towns was 142 days after the epidemic arrived in Belfast (range day 1 – day 291; median day 137; IQR day 106.8 – day 164.2). The average epidemic day of last reported cholera case for each infected town was day 248 (range day 89 – day 573; median day 238; IQR day 176.8 – day 298.5). The average number of days from first reported cholera case to last for the infected towns was 106 days (range 2 days – 491 days; median 75 days; IQR 46.7 days – 140 days). The pattern of this spread was illustrated through an isopleth map showing the epidemic day of the first case for each town in Figure 14. This map shows cholera cases first being reported on the North-Eastern coast of Ireland, with some sporadic spread to follow on the Central-Eastern coast. It is not until epidemic day 120 that the North-Eastern outbreak connects with the South-Western outbreak, suggesting the possibility of a separate arrival of cholera from the South-Western coast.

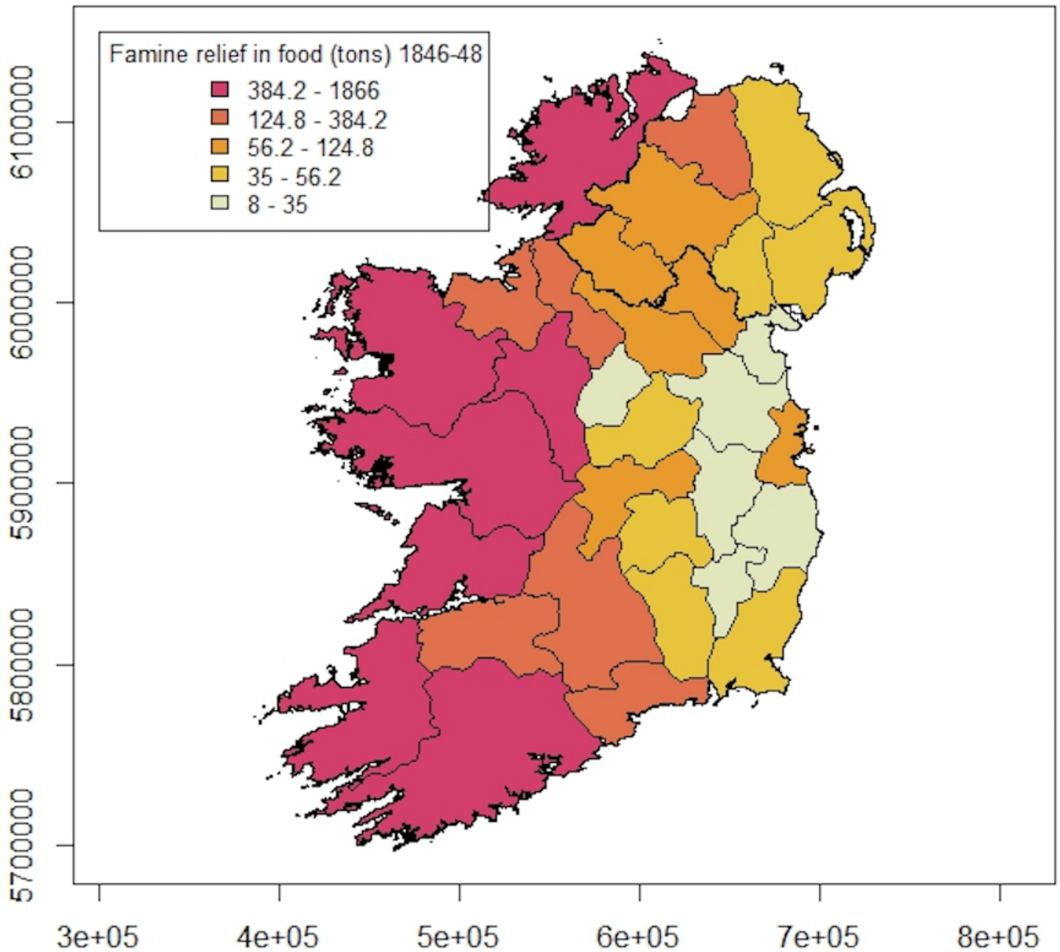

Figure 8. Total Quaker famine relief in food (tons) by county in Ireland (1846 – 1848).

Analysis of the cholera epidemic at county and kingdom levels

Spatial

The estimated attack rate for all of Ireland was 697 cases per 100,000 people (95% CI 691.07 – 703.83) and the estimated case fatality rate for all of Ireland was 42% (95% CI 41.84 – 42.74).

Risk maps

Figure 15 shows the cholera attack rate per 100,000 people, by county in Ireland, 1848 to 1850. Isopleth maps illustrate the pattern of cholera attack rates and mortality rates within the total reported population, pointing out high risk of cholera infections and mortality in the Central-Western coast of Ireland (Figures 16 and 17). An isopleth map of cholera case fatality rates illustrates a different pattern, with high case fatalities in Central Ireland, the South-Western coast, and the South-Central coast (Figure 18).

Cluster detection

Figures 19 and 20 visualize the location of significant cholera attack rate clusters and cholera case fatality rate clusters, respectively. The most likely cholera attack rate cluster is in Galway county (Relative Risk (RR) = 3.42; 95% CI 2.61 – 4.49; $p < 0.001$). Further secondary cholera attack rate clusters are Antrim county (RR = 1.46; 95% CI 1.34 – 1.59; $p < 0.001$), King's county (RR = 1.65; 95% CI 1.48 – 1.84; $p < 0.001$), and Waterford county, Kilkenny county, Tipperary county, Wexford county, Cork county and

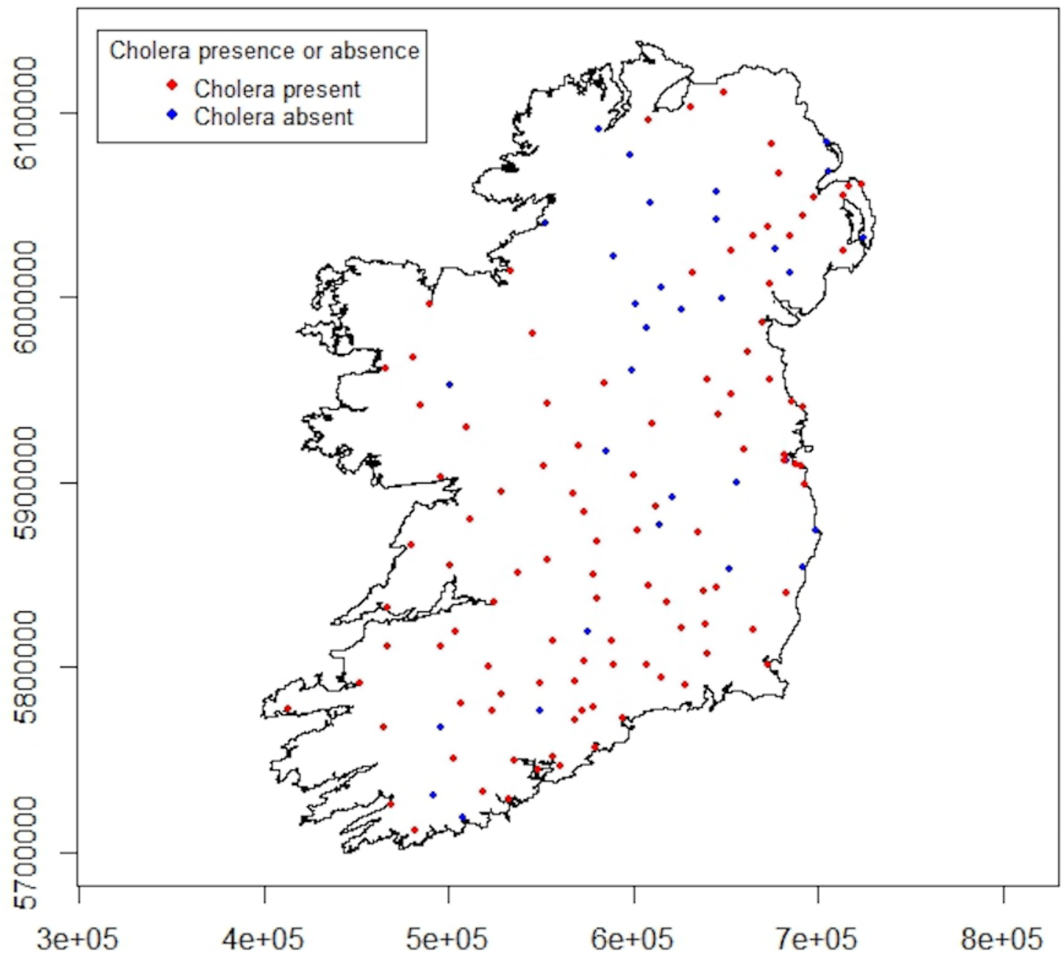

Figure 9. Presence and absence of reported cholera cases by major cholera town (1848 – 1850).

Carlow county (RR = 1.14; 95% CI 1.11 – 1.17; $p < 0.001$). The most likely cholera case fatality cluster includes King's county, Queen's county, Westmeath county, Kildare county, Longford county, Tipperary county, Roscommon county, Kilkenny county, Carlow county, and Meath county (RR = 1.15; 95% CI 1.10 – 1.20; $p < 0.001$). Further secondary cholera case fatality clusters include Leitrim county and Sligo county (RR = 1.24; 95% CI 1.13 – 1.36; $p < 0.001$), Kerry county (RR = 1.31; 95% CI 1.15 – 1.48; $p < 0.001$), and Waterford county (RR = 1.13; 95% CI 1.04 – 1.22; $p = 0.0019$). Of the eighteen different counties that make up the cholera attack rate and case fatality rate clusters, a total of five counties were part of both: the cholera attack rate clusters and the case fatality rate clusters.

Temporal

Figure 21 shows the epidemic curve for 1848 – 1850. The main peak of the epidemic was in the spring of 1849, peaking in the month of May with 11,129 new cases. A smaller peak followed in August 1849 with 4,905 cases, before dropping significantly in October 1849. The final case was reported on June 29, 1850.

Modelling cholera attack and case fatality rates

Table 3 and 4 show the results of the spatial generalized linear mixed models of cholera attack rates and cholera case fatality rates, respectively. Percent persons illiterate was removed because an association of r = 0.71 (95% CI 0.48 - 0.85) with percent families in 4th class houses was measured, and percent families

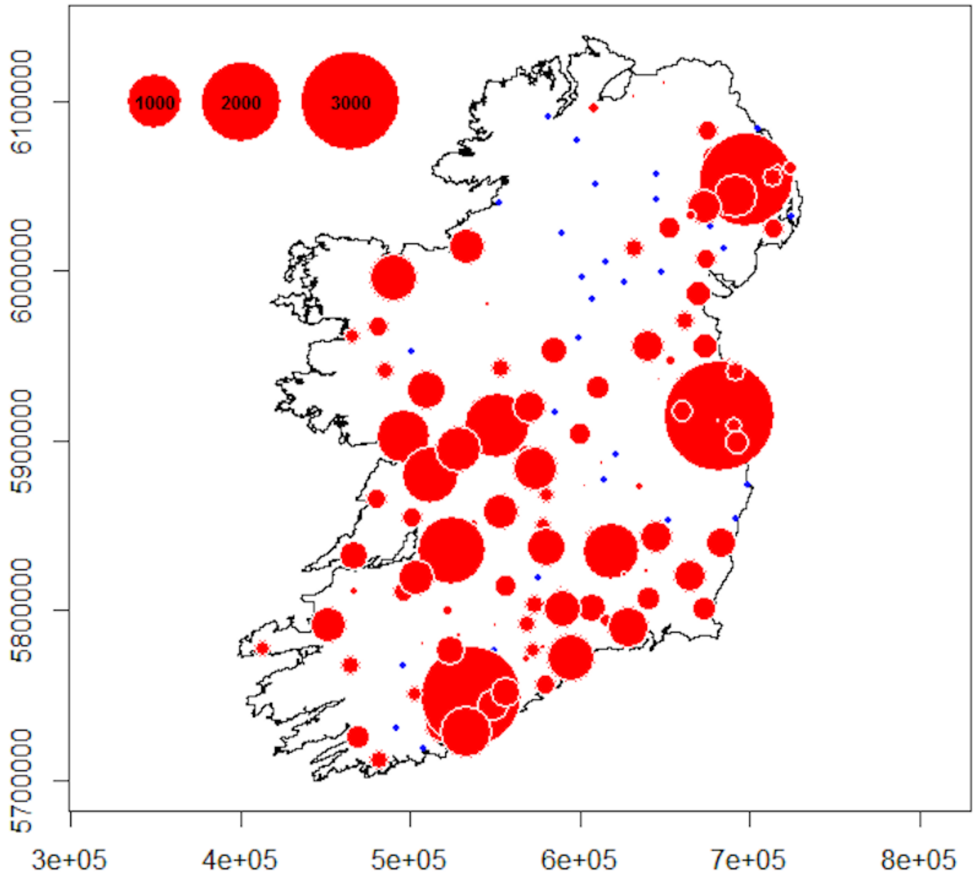

Figure 10. Total number of reported cholera cases per major town in Ireland (1846 – 1850).

in 4th class houses was judged to be a more direct measurement (Table 5). Normal quantile-quantile plots of deviance residuals and observed versus predicted plots for the final cholera attack rate models revealed no lack of fit, with some homoskedasticity of the residuals. The range of the semivariogram for the attack rate model was 28.9 kilometres, indicating reasonable spatial dependence. Normal quantile-quantile plots of deviance residuals and observed versus predicted plots for the case fatality rate models revealed no lack of fit. The range of the semivariogram for the case fatality rate models was 28.3 kilometres, indicating reasonable spatial dependence. No interaction terms were significant.

Cholera attack rates were most associated with population health, number of families per 100 houses, population density, and poverty, although none of the associations were significant (Table 3).

Case fatality rates were most associated with number of families per 100 houses, population density, poverty, and famine. The risk of case fatality decreases by 0.98 times for every additional family per 100 houses ($p = 0.008$), when controlling for the other covariates. The case fatality rate of cholera was not significantly associated with the other model covariates (Table 4). Smoothed attack rates were also not significantly associated with smoothed case fatality rates ($p = 0.52$).

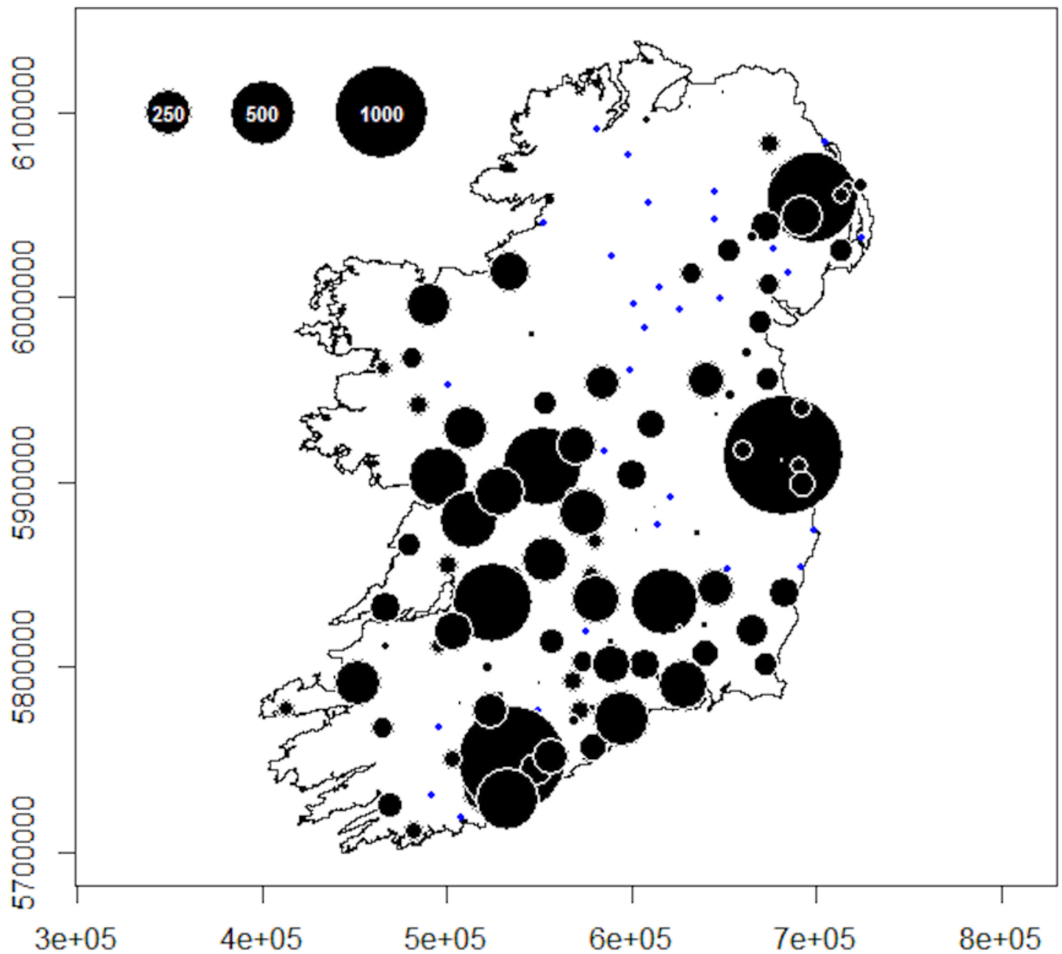

Figure 11. Total number of reported cholera deaths per major town in Ireland (1848 – 1850).

DISCUSSION

Those regions where a cholera case was least likely to survive did not correspond to where a cholera case was most likely to occur. Of the 32 counties of Ireland, 18 were identified as part of clusters of either elevated attack rates or case fatality rates. However, these clusters overlapped only by 5 counties and covered otherwise different regions of the study area. The majority of counties that experienced higher than expected attack rates did not have higher than expected case fatality rates. Similarly, the majority of counties that experienced higher than expected case fatality rates did not have higher than expected attack rates. Therefore, the factors that influenced the risk of contracting cholera during the epidemic of Ireland from 1848 – 1850 may not have been the same as the factors that influenced the risk of dying from cholera once infected.

The cholera epidemic of Ireland began on the North-Eastern coast, continuing on the Central-Eastern coast, with a possible separate arrival via the South-Western coast. The arrival of cholera by coast is consistent with other cholera outbreaks; as John Snow reported in 1855, cholera "always appears first at a sea-port" (Snow, 1855). Cholera was reported in a majority of the major towns of Ireland, although the attack rates, mortality rates, and case fatalities varied. This epidemic had significant clusters where counties had higher than expected attack rates or higher than expected case fatality rates when compared to other counties.

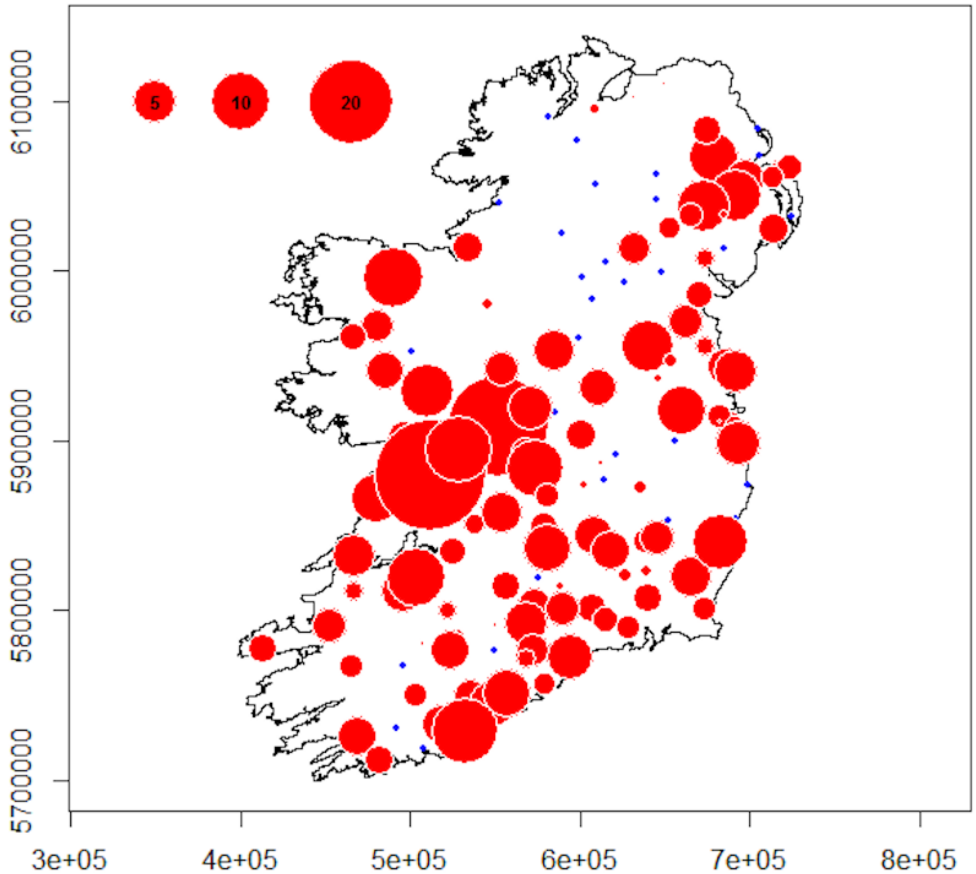

Figure 12. Cholera attack rates per major town in Ireland: percent of population that contracted cholera (1848 – 1850).

Risk of contracting cholera during the cholera epidemic of Ireland from 1848 – 1850 was not significantly associated with the demographic and famine variables included in this study. John Snow came to a similar conclusion in 1855 following cholera outbreaks in London; he found that class did not appear to impact who contracted cholera, but instead cholera infection "appears to have fallen pretty equally amongst all classes, in proportion to their numbers" (Snow, 1855). He also found that housing "quality" and occupation did not impact risk of cholera (Snow, 1855). Other factors contributed to the spatial variations in cholera attack rates and the higher than expected cholera attack rate clusters apparent in this cholera epidemic. For Snow, it was not demographics but the source of water that explained where cholera cases occurred, where "each epidemic of cholera in London has borne a strict relation to the nature of the water-supply of its different districts" (Snow, 1855), and the same was likely true for this epidemic in Ireland.

However, demographic variables were associated with the cholera case fatality rates during this epidemic. The risk of dying from cholera once infected decreased slightly for every additional family living per 100 houses, which was used here as an indicator for the known risk factor of overcrowding. Both the Commissioners and the Medical Officers repeatedly identified Ireland's rural population of being at greater risk for death by cholera (Rep, 1852). Population density was not found to be significantly

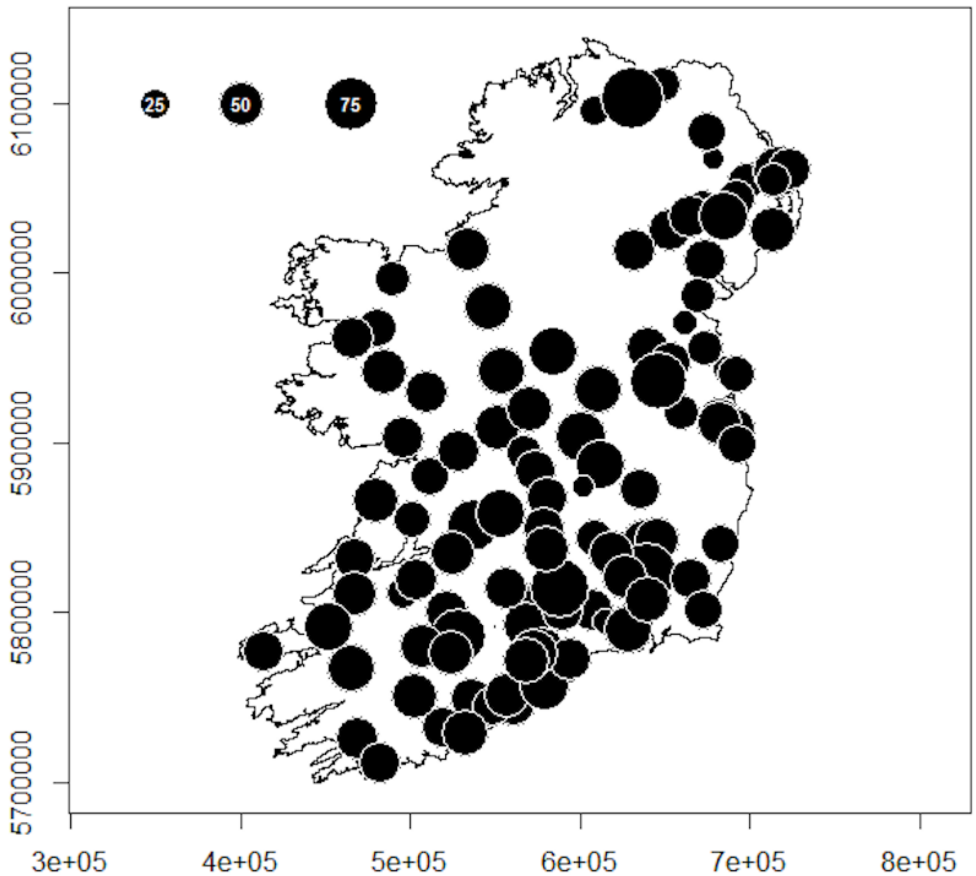

Figure 13. Cholera case fatality rates per major town in Ireland: percent of cholera cases that died of cholera (1848 – 1850).

associated with attack rates or case fatality, but those infected by cholera in counties with more families living per 100 houses were less likely to die of their infection than those in less dense housing situations. If the densely occupied houses were also more urban, the Commissioners and Medical Officers may have been correct in assuming that the rural population was more at risk of dying from cholera than those in urban areas. The Medical Officers in the Commissioner's report were almost unanimous in citing the famine and poverty as the main reasons for the abnormally high case fatality of their patients during this period (Rep, 1852), and previous research has associated population density (Rep, 1852) with cholera, but this study did not find these associations significant.

Limitations

The data studied here only represent the reported cholera cases and deaths, i.e. those reported to Ireland's Central Board of Health. The Commissioners discussed the issue of under-reporting, saying "the deaths of all who die in hospital are registered; but there is no record of those who perish on the road sides, or in their own abodes" (Rep, 1852). Limiting the study to only reported hospitalized cases could have introduced bias if these data were systematically different; hospitalized patients may be worse clinically and more likely to die, or alternatively only healthy patients survived long trips to the hospital. There may have

Variable	Value	SE	Degrees of freedom	t-value	p-value	IRR
(Intercept)	-17.796	7.353	27	-2.420	0.0225	0.000
Percent with health conditions	-0.215	0.823	27	-0.261	0.7963	0.807
No. families per 100 houses	0.116	0.074	27	1.564	0.1294	1.123
No. persons per km^2	0.002	0.011	27	0.215	0.8311	1.002
Percent families in 4th class houses	0.070	0.052	27	1.343	0.1903	1.073

Table 3. Spatial GLMM model for attack rates by demographic covariates. Non-significant covariates were not removed if they changed remaining covariate parameter values by greater than 10%.

Variable	Value	SE	Degrees of freedom	t-value	p-value	IRR
(Intercept)	2.506	0.557	27	4.501	0.0001	12.259
No. families per 100 houses	-0.016	0.005	27	-2.840	0.0085	0.985
No. persons per km^2	0.002	0.001	27	1.753	0.0910	1.002
Percent families in 4th class houses	-0.008	0.004	27	-1.734	0.0943	0.992
Famine relief in food (tons)	0.000	0.000	27	1.491	0.1476	1.000

Table 4. Spatial GLMM for cholera case fatality rates by demographic covariates.

been underreporting at the individual and hospital level. At the individual level, underreporting would be expected due to the social stigma associated with illness and a distrust of the overcrowded hospitals (Rep, 1852). At the hospital level, it is possible that underreporting could have occurred due to fear of quarantine and trade restrictions. Bias from underreporting could have occurred; for example hospital cases were more likely to be reported if they died. Additionally, during this period there was no way to clinically distinguish cholera from other diarrhoea (Chan et al., 2013). Without clinical confirmation of cases, misclassification must be assumed for these data.

Additionally, studying a historical epidemic requires making the best of the available data, which is often sparse. For example, many of the covariates used to evaluate the demographics were proxies of the actual covariates of interest, which limits the conclusions that can be drawn. These covariates were only available at the county level, which limited the sample size to 32 and lowered the power of the study. All p-values are considered exploratory p-values. The study is an ecological study which constrains the interpretation of associations between the covariates and the epidemic variables; to avoid an ecological fallacy, inferences could not be made about the individuals in Ireland during the epidemic as the data is measured at the town, county, or kingdom level.

The census of Ireland most relevant to the epidemic was conducted in 1851. Therefore, demographic data used from this census are from after the cholera epidemic because demographic data were not available for the epidemic period itself. Using demographic data from the end of the epidemic increases the risk of survivor bias. Similarly, the famine relief data were only available for 1846 – 1848, which did not coincide exactly with the cholera epidemic.

Strengths

Despite the limitations intrinsic with historical data, the *Report of the Commissioners of Health, Ireland, on the epidemics of 1846 to 1850* describes the system of data collection and the situation in Ireland in great detail, allowing assessment of the quality and completeness of the report (Rep, 1852). Additionally, the *Census of Ireland for the Year 1851* (Cen, 1856) is considered one of the most thorough national censuses ever conducted (Froggatt, 1965). Therefore, there is confidence that these historical data could be considered more complete and systematic than for most other historical epidemics.

CONCLUSIONS

During the cholera epidemic of Ireland of 1848 – 1850, the majority of counties that experienced higher than expected attack rates did not have higher than expected case fatality rates. Similarly, the majority of counties that experienced higher than expected case fatality rates did not have higher than expected attack rates. Therefore, the factors that influenced whether people contracted cholera were not the same as the factors that influenced their risk of dying of cholera once infected. During this epidemic, people of all

Variable	Value	SE	Degrees of freedom	t-value	p-value	IRR
(Intercept)	2.506	0.557	27	4.501	0.0001	12.259
No. families per 100 houses	-0.016	0.005	27	-2.840	0.0085	0.985
No. persons per km^2	0.002	0.001	27	1.753	0.0910	1.002
Percent families in 4th class houses	-0.008	0.004	27	-1.734	0.0943	0.992
Famine relief in food (tons)	0.000	0.000	27	1.491	0.1476	1.000

Table 5. Spearman rank-based measures of association of county level dependent and explanatory variables.

education levels, health conditions, housing conditions, and poverty levels, were at risk of contracting cholera. This outbreak provides insight into the natural history of cholera before there were effective public health interventions or public understanding of transmission. Cholera infection spread widely to the majority of major towns, no matter their prosperity or demographics. Without interventions, everyone is at risk for cholera infection.

Although everyone was at risk of contracting cholera during this outbreak, some were more likely to die of their infection. The Commissioners and the Medical Officers of the time may have been correct in their assessment that the health infrastructure of rural Ireland was inadequate, as houses populated by fewer families were more at risk for dying of cholera than those living in the more densely populated houses. This study provides evidence that people that were housed more densely had better chance of surviving their cholera infections. Before there were public health interventions to prevent cholera infection, access to health infrastructure can be the difference between life and death. This finding may emphasize the importance of health care provisions in disaster zones such as emergency care for populations affected by earthquakes and hurricanes.

Competing interests

None of the authors has any competing interests.

Authors' contributions

ACS: Project leadership, data collection, design and execution of quantitative analysis, interpretation of results; drafted the manuscript. OB: Project leadership, data collection, design of quantitative analysis and interpretation of results. EK and JS: supported in the design of the study, interpretation the results and writing of the manuscript. All authors have revised and contributed significantly to the final manuscript. All authors have read and approved the final manuscript.

ACKNOWLEDGEMENTS

ACS was supported by the Laforet Research Assistantship through the Ontario Veterinary College.

REFERENCES

(1852). Extracts from transactions of the central relief committee of the society of friends during the famine in ireland in 1846 and 1847.

(1852). Report of the commissioners of health, ireland, on the epidemics of 1846 to 1850: presented to both houses of parliament by command of her majesty.

(1856). The census of ireland for the year 1851: presented to both houses of parliament by command of her majesty.

(2014). Cholera: Fact sheet.

Arnold, D. (1993). Social crisis and epidemic disease in the famines of nineteenth-century india. *Social history of medicine*, 6(3):385–404.

Berke, O. (2004). Exploratory disease mapping: kriging the spatial risk function from regional count data. *International Journal of Health Geographics*, 3(1):18.

Bingham, P., Verlander, N., and Cheal, M. (2004). John snow, william farr and the 1849 outbreak of cholera that affected london: a reworking of the data highlights the importance of the water supply. *Public health*, 118(6):387–394.

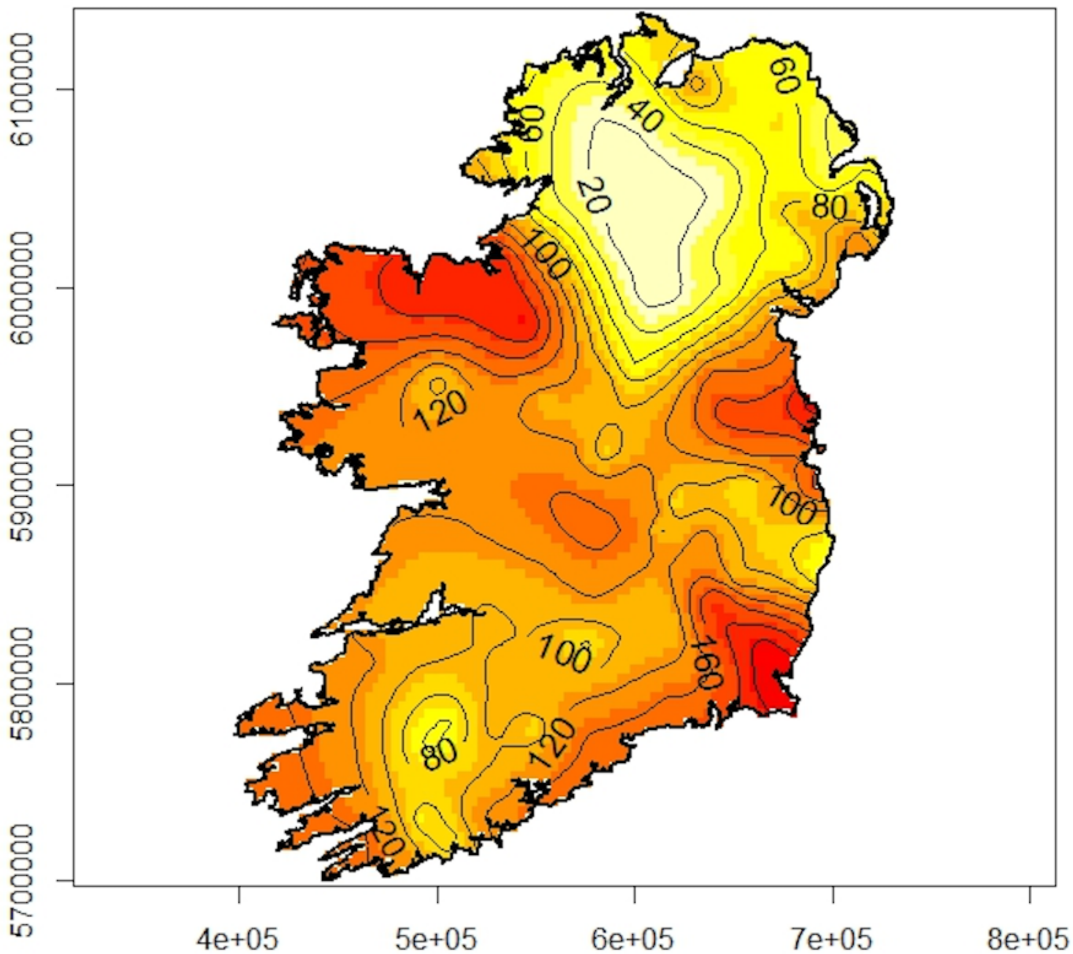

Figure 14. Risk map showing the epidemic day of first reported cholera case for major towns of Ireland (1848 – 1850).

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421):9–25.

Chan, C. H., Tuite, A. R., and Fisman, D. N. (2013). Historical epidemiology of the second cholera pandemic: Relevance to present day disease dynamics. *PloS one*, 8(8):e72498.

Cressie, N. and Hawkins, D. M. (1980). Robust estimation of the variogram: I. *Journal of the International Association for Mathematical Geology*, 12(2):115–125.

Dohoo, I., Martin, S., et al. (2003). *Veterinary epidemiologic research*. Number V413 DOHv.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., and Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1):27–46.

Dormann, C. F., M McPherson, J., B Araújo, M., Bivand, R., Bolliger, J., Carl, G., G Davies, R., Hirzel, A., Jetz, W., and Daniel Kissling, W. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5):609–628.

Froggatt, P. (1965). Sir william wilde and the 1851 census of ireland. *Medical history*, 9(04):302–327.

Hjalmars, U., Kulldorff, M., Gustafsson, G., and Nagarwalla, N. (1996). Childhood leukaemia in sweden: using gis and a spatial scan statistic for cluster detection. *Statistics in Medicine*, 15(7-9):707–15.

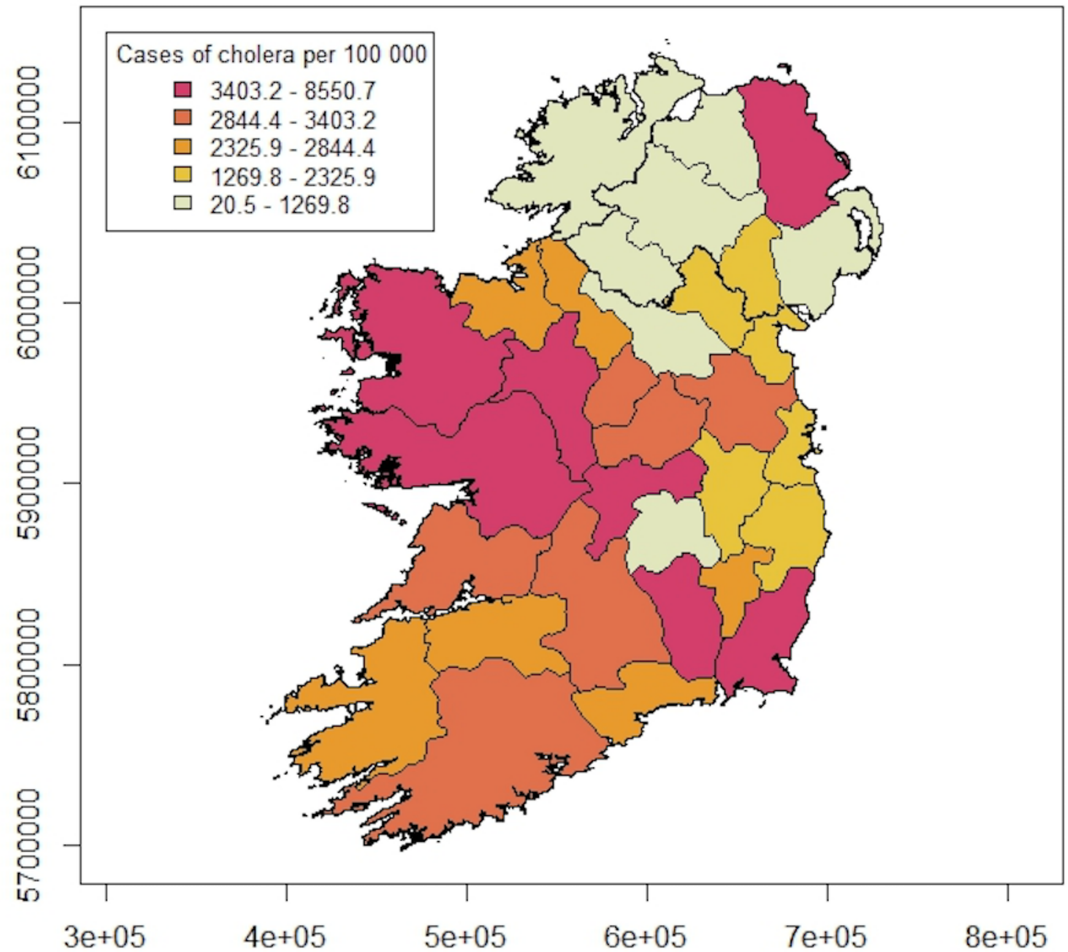

Figure 15. Number of reported cholera cases per 100 000 people in Ireland (1848 – 1850).

Jutla, A., Whitcombe, E., Hasan, N., Haley, B., Akanda, A., Huq, A., Alam, M., Sack, R. B., and Colwell, R. (2013). Environmental factors influencing epidemic cholera. *The American journal of tropical medicine and hygiene*, 89(3):597–607.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics-Theory and methods*, 26(6):1481–1496.

Kuo, C.-L. and Fukui, H. (2007). Geographical structures and the cholera epidemic in modern japan: Fukushima prefecture in 1882 and 1895. *International journal of health geographics*, 6(1):25.

Nolan, W. (2014). Geography of ireland: Description of the physical and human geography of ireland.

Snow, J. (1855). *On the mode of communication of cholera*. John Churchill.

Tauxe, R. V., Holmberg, S. D., Dodin, A., Wells, J. V., and Blake, P. A. (1988). Epidemic cholera in mali: high mortality and multiple routes of transmission in a famine area. *Epidemiology and infection*, 100(02):279–289.

Young, H. and Jaspars, S. (1995). Nutrition, disease and death in times of famine. *Disasters*, 19(2):94–109.

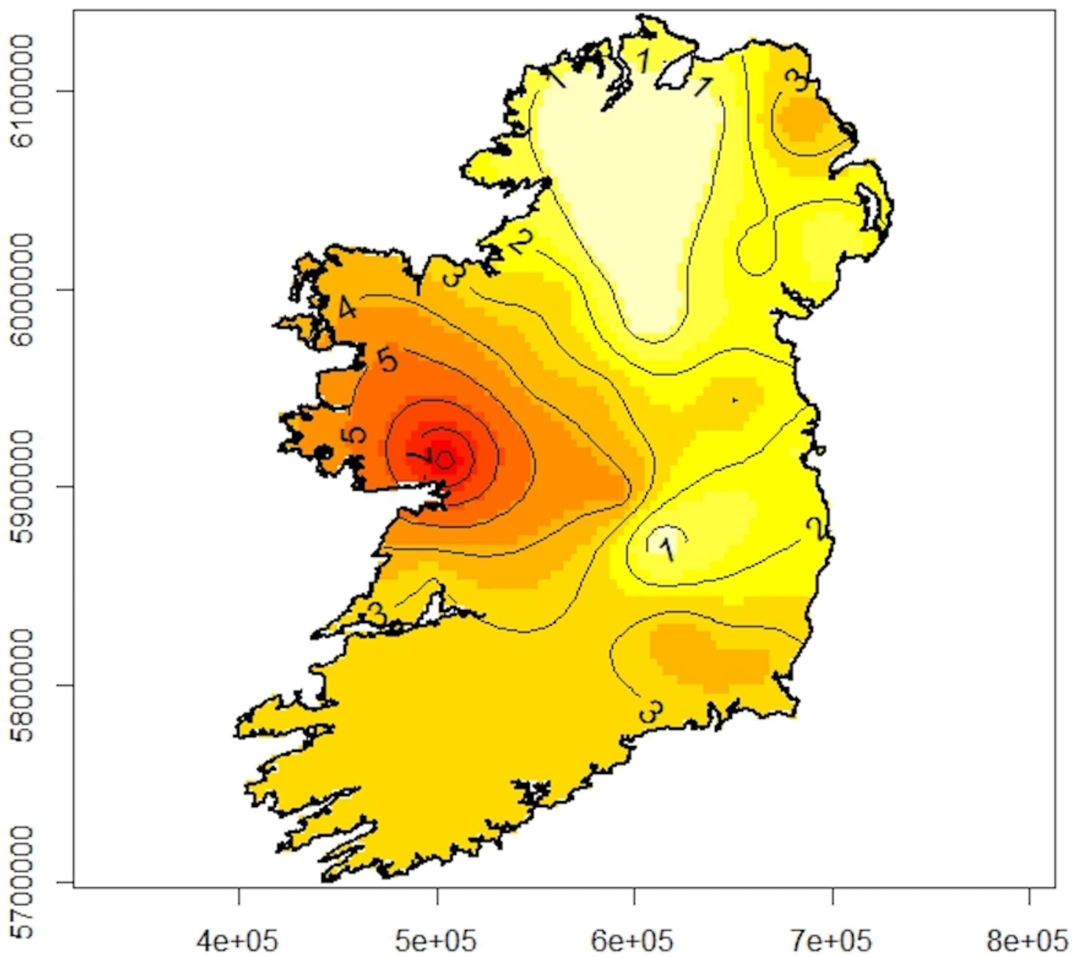

Figure 16. Risk map of the attack rates of cholera in Ireland: percent of population that contracted cholera (1848 – 1850).

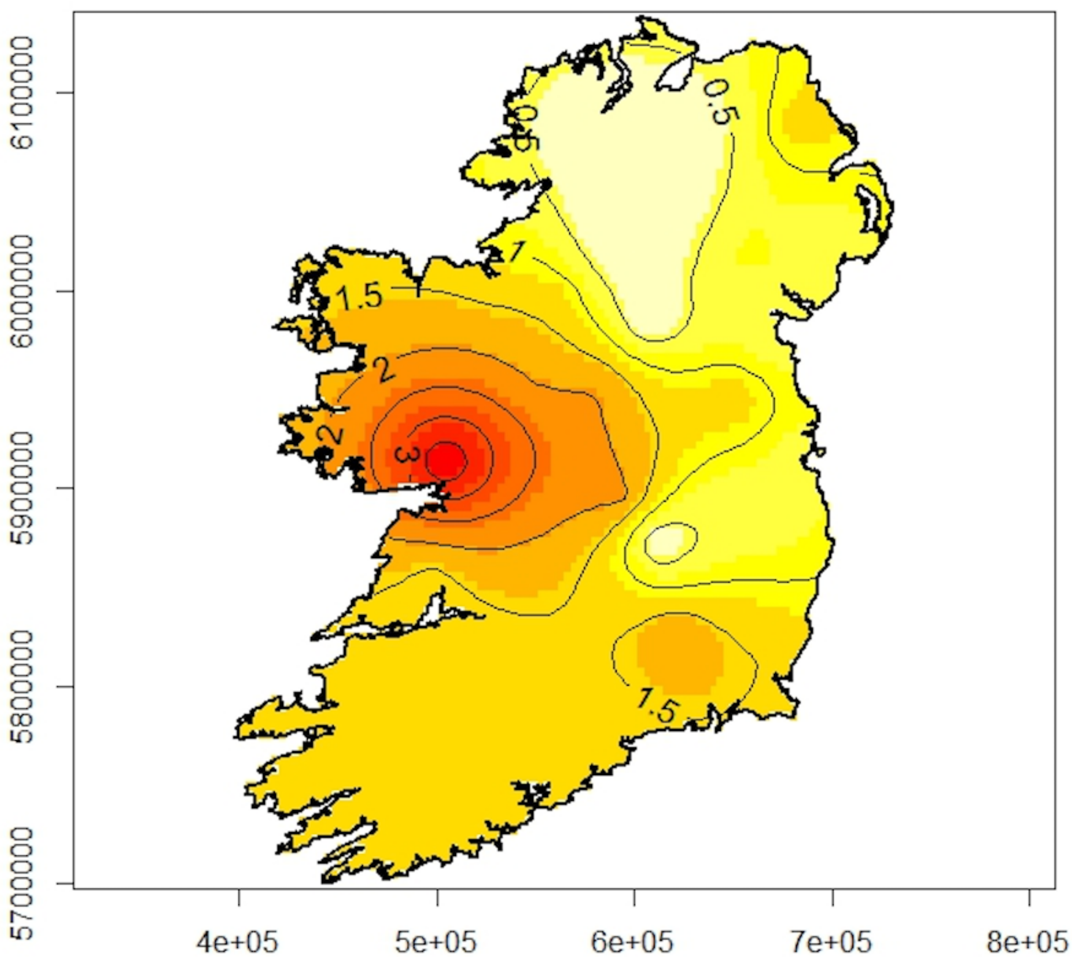

Figure 17. Risk map of the mortality rates of cholera in Ireland: percent of population that died of cholera (1848 – 1850).

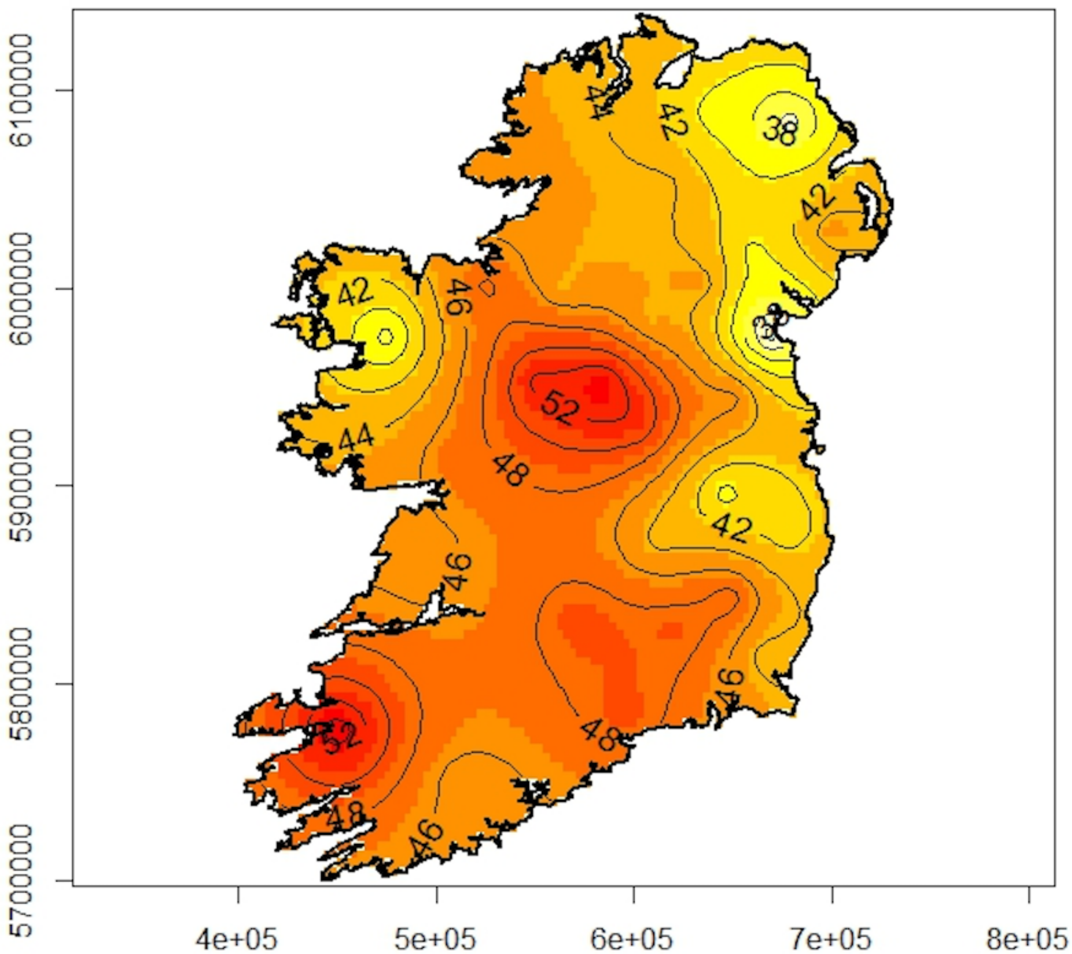

Figure 18. Risk map of the case fatality rates of cholera in Ireland: percent of cholera cases that died of cholera (1848 – 1850).

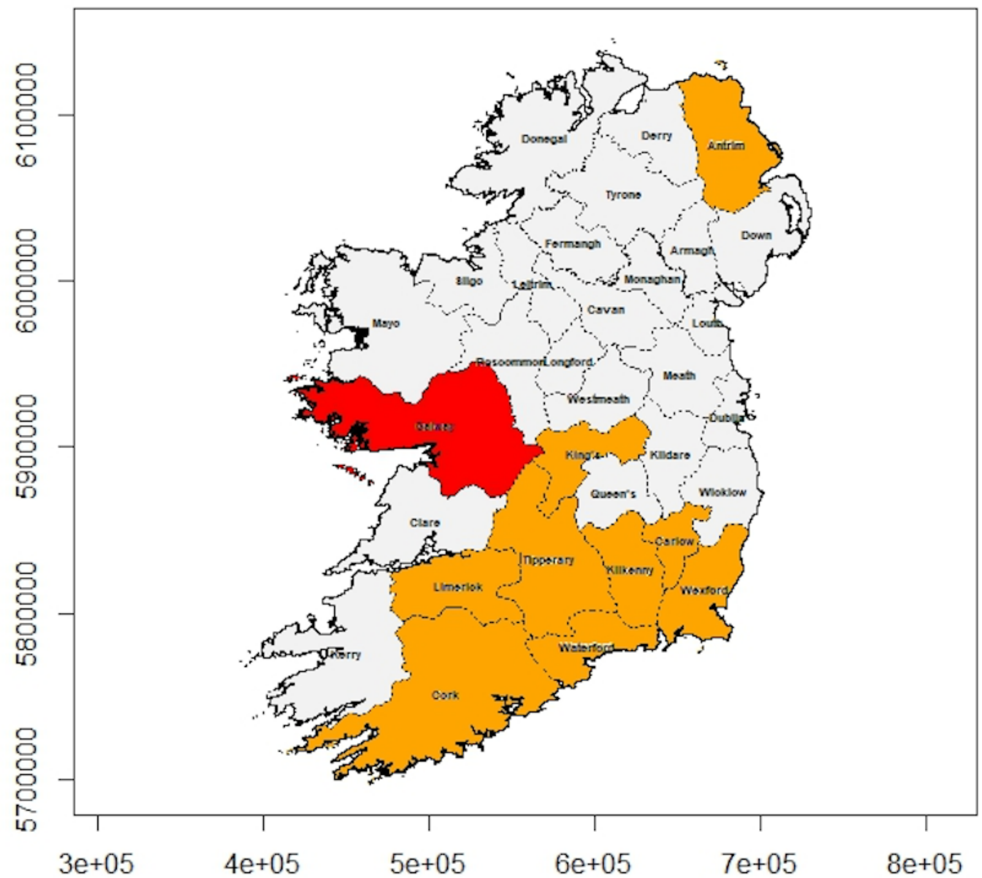

Figure 19. Most likely cholera attack rate clusters and secondary clusters in Ireland (1848 – 1850).

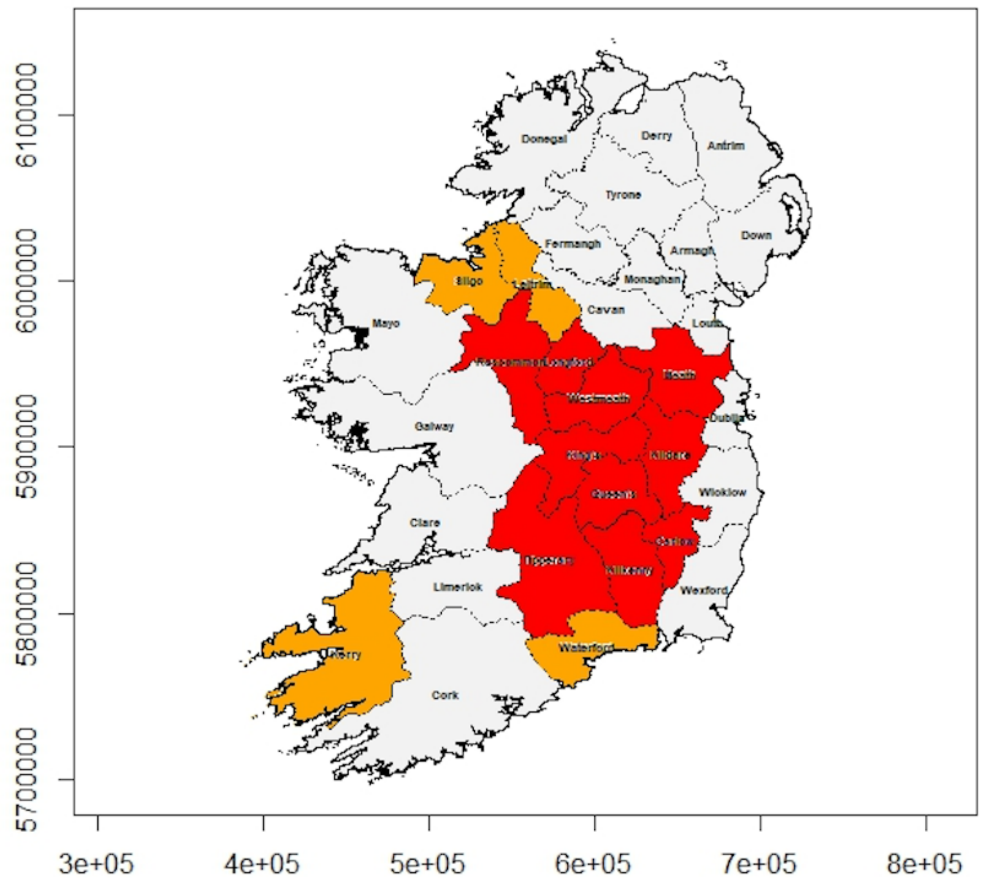

Figure 20. Most likely cholera case fatality rate clusters and secondary clusters in Ireland (1848 – 1850).

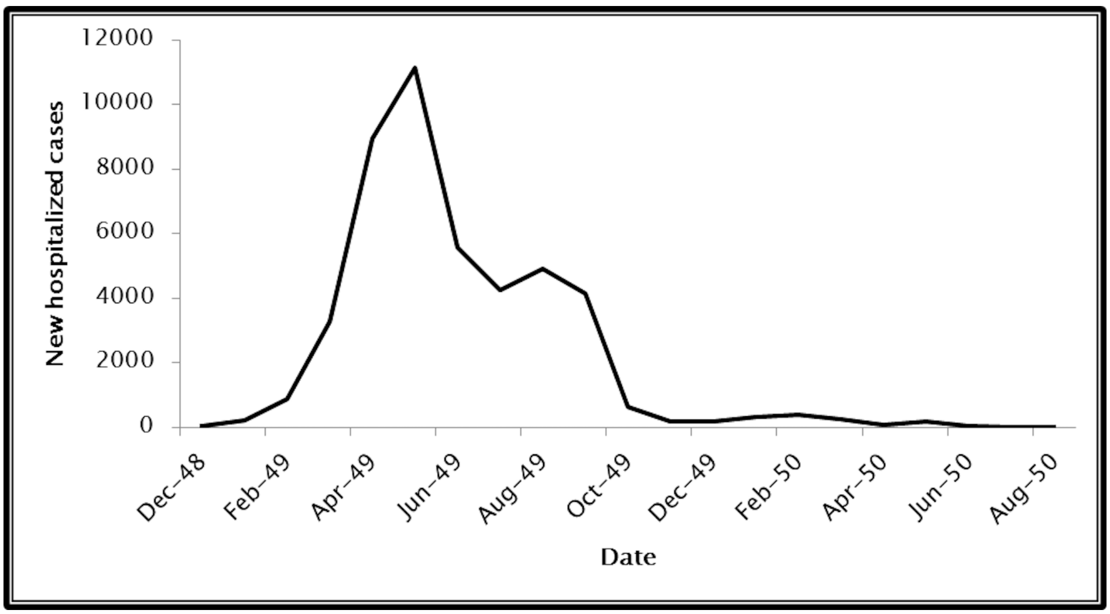

Figure 21. Epidemic curve of total monthly reported new cholera cases in Ireland (1848 – 1850).