Spatial comparison of herald and main waves in London's nineteenth-century cholera epidemics

Walter Peterson

Captain, US Navy (Retired), Stafford, VA, USA

Abstract

Nineteenth-century London experienced four extraordinarily severe summertime cholera epidemics. Three were preceded by less severe non-summer outbreaks. Twenty-first-century research hypothesizes them as herald waves of potentially new cholera strains. This study examined the geographical characteristics of these herald waves and compared them to their subsequent main waves to determine if there was a geographical component to the significant difference in wave severity. Cholera mortality data for London's parishes and registration districts were extracted from contemporaneous records. The data were normalized and scaled. Each epidemic wave was divided into two segments for analysis. A Spearman's rank correlation was used to assess the relationship between a herald and its subsequent main wave. Geospatial analytical tools were used to determine and display each segment's geographic distribution pattern using autocorrelation techniques to determine its central point. Results show that the herald wave of each epidemic shared characteristics similar to its following main wave. Central-point locations were similar and Spearman's rank coefficients showed high degrees of correlation. Autocorrelation results were similar, with one exception reflecting an appalling anomalous cholera outbreak at an institution for children. Because of the demonstrated similarity of each epidemic's herald and main waves, this study did not detect a spatial characteristic that could explain the observed difference in severity between the studied heralds and mains.

Introduction

In 2010, an eclectic team of Canadian and American mathematicians, statisticians, and epidemiologists studied the temporal characteristics of London's cholera mortality during the years 1824 to 1901 (Tien *et al.*, 2011). The team determined that cholera deaths were seasonal and usually occurred in the summer months. Their report stated: '…*the only non-summer outbreaks occurred in the spring of 1832, the autumn of 1848 and the winter of 1853 and extraordinarily severe [subsequent] summer outbreaks occurred in 1832, 1849, 1854…*' The report further noted that '*The non-summer outbreaks of 1832, 1848 and 1853 appear to have been herald waves of newly invading cholera strains.*' (Tien *et al.*, 2011). The notion of herald waves had not been previously associated with cholera outbreaks.

The three herald waves identified during the nineteenth century are unique among London's historical mortality records. Also noteworthy is that herald waves associated with contemporaneous cholera epidemics have not been seen in other major cities in the United Kingdom (UK) or in Europe. Although such waves have been observed in influenza epidemics, a review of the relevant literature shows that studies focused on the wave temporal mortality and did not include spatial comparisons between herald and main waves. This study examined the geographic characteristics of the three cholera epidemics that started in 1832, 1848 and 1853, respectively. The aim was to determine if, in addition to the observed significant severity differences, spatial differences also existed between the heralds and their following main waves that might have been in part responsible for the differences in severity.

Materials and methods

Background

The 157 parishes within London's 1832 Bills of Mortality (LBoM) define the study area for the 1832 epidemic. The UK Data Service Study 4828 (Burton and Southall, 2004) provides the geographical information systems (GIS) parameters to build parish geometries as they existed in 1832. The Births and Deaths Registration Act of 1836 replaced the Bills of Mortality as the venue for registering and reporting mortality. The act established 36 Registration Districts and 135 Registration sub-districts in the London metropolitan area for reporting purposes and these jurisdictions form the study area for the 1848/49 and 1853/54 epi-

demics. The UK Data Service Study 852232 (Satchell *et al.*, 2018) provides the GIS parameters that define these districts.

Study area

Figure 1 provides geographic context and an overview comparison of the area covered by the 1832 LBoM. The area of London within the original Roman walls of the ancient city is referred to as 'within the walls' in the LBoM. Today it is often simply referred to as The City or The Square Mile. In the Registration District paradigm, it is known as London City, the outline of which is shown for reference purposes. It makes up only a small part of Metropolitan London. During the 1832 epidemic it contained 97 of the 157 parishes in the LBoM. During the later epidemics it was a single registration district.

Data sources

The LBoM for 1832 (Worshipful Company of Parish Clerks, 1832) provided parish weekly cholera mortality data for the 1832 epidemic. The Metropolitan Sanitary Commission First Report of the Commissioners (1848) also provided a summary of 1832 mortality; however, it is grouped by the later registration district paradigm. The Commission data are not sorted by week but presented as registration district totals for the entire epidemic. The Registrar General's Weekly Reports (RGWR) of the Weekly Return of Births and Deaths in London for the 1848/49 and 1853/54 epidemics provided cholera mortality data, which are bound together in annual volumes in London's Wellcome Library, with some annual volumes available online at the Hathi Trust Digital Library (https://www.hathitrust.org/).

Notes on the data

Shortcomings exist within the LBoM; principally, they only record burials (not deaths) within the Anglican parish burial grounds and exclude other protestant denominations, Roman Catholics, Jews and '*dissenters*.' Commentators have also noted other problems with the Bills and, particularly relevant to this study, frequently cite that parish reporting became erratic in the early part of the nineteenth-century. For example, the parish of St. George, Hanover Square, entirely stopped submitting returns in 1823 (Maplestone, 2011). One effect of these inconsistencies is easily confirmed by comparing the deaths reported in the 1832 LBoM for the 97 parishes '*within the walls*' with the number reported in the Metropolitan Sanitation Commission Report for the same area; the LBoM recorded 152 cholera deaths, while the Commission reported 359. This and other deficiencies do not, however, completely preclude the value of using the LBoM as a relative measure to compare the 1832 herald wave with the 1832 main wave. As noted, there are 97 parishes located '*within the walls*.' Their combined area is on the same order of magnitude as one of the 60 individual parishes 'outside' the walls. For processing purposes, the 97 have been combined together into a single parish unit. This then gives a total of 61 parishes. The previously noted deficiencies in parish reporting leave six without a non-zero pre-epidemic average total mortality. Consequently, these six were merged with an adjoining parish to prevent division by zero during the normalization process. The cholera deaths in the '*deleted*' six parishes are included in the totals for the newly created 'merged' parishes. All 3,200 cholera deaths recorded in the LBoM remain in the redrawn 55 parishes and are included in all 1832 epidemic calculations. The RGWRs for some years contain total, all-cause mortality data for preceding years. These data form the basis for calculating an average pre-epidemic total mortality used for normaliza-

tion purposes. Occasionally from one year to the next, the reported preceding year's data are inconsistent. For example, the average of total mortality reported for London's registration districts for the years 1850, 1851 and 1852 in the RGWR for 1853 differs from that reported in the RGWR for 1854 by 55 deaths. This difference is insignificant when compared with the over 53,000 average mortalities for the three years. The data from the 1853 RGWR are the most often repeated in subsequent annual summaries issued by the Registrar General and are the data used in this study.

Data processing

The software package ArcGIS Pro (2.7.0) (ESRI, Redlands, CA, USA) was used to spatially analyse and display the mortality data.

Approach

In the broadest of terms, this study consists of three individual and independent studies. Each study spatially compared a herald wave with its immediately following main wave in one of London's 1832, 1848/49 or 1853/4 cholera epidemics. Geographic distributions, weighted mean centres, autocorrelation patterns and Spearman's correlation coefficients formed the base for wave comparisons, which focused at two key temporal points, dividing each wave into two segments: the build-up and the end. Since most epidemics approximate exponential growth during their initial phase (Ma *et al.*, 2014), this study used the end of the exponential growth period as one point of wave comparison. Weekly cholera mortality ceasing or returning to an endemic level defined the other point of comparison. The cumulative parish/registration district cholera mortality totals at these two key points formed the basis for wave comparisons. Either the LBoM or the RGWR, as appropriate, provided the location information for weekly death counts. Normalizing the death count data provided the basis for comparing herald with main parish/registration district cholera mortality. In the absence of reliable mid-census population data, the normalization process used the average total, all-cause mortality for three preceding non-epidemic years as the normalizing parameter. The process concluded with a 0 to 1 scaling of the normalized results.
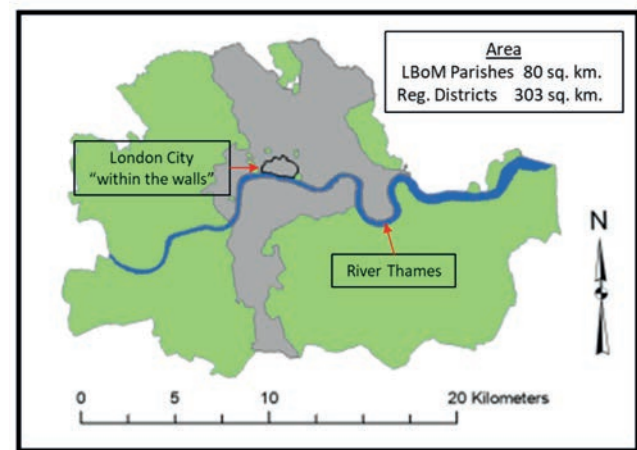

Figure 1. Areas of interest for the study of London's 19th century cholera epidemics. The area in grey represents the area contained within the 1832 London Bills of Mortality, while this area plus the green one represents the coverage of the Registration District paradigm instituted by the Births and Deaths Registration Act of 1836.

Mapped wave scaled data provided for visual inspection and comparison. Analyses of quantile-quantile plots confirmed that cholera deaths among the parish/registration districts were not normally distributed. Consequently, a Spearman's rank correlation coefficient calculation compared the distribution of parish/district cholera deaths between heralds and mains at the key points.

Multiplying the x and y coordinates of each parish/registration district's geographic centre with a weighting factor representing the feature's mortality, then summing the x and y products and dividing by the segment's total cholera mortality produced a weighted mean centre. This can be thought of as the segment/wave's centre of '*mortality*' and it would then be useful for comparison with the other segments/waves. Moran's Index (Moran's I) (Moran, 1950) was used to determine a global autocorrelation coefficient when comparing the mortality between neighbouring parishes/registration districts at the key points. The algorithm evaluated whether the mortality pattern was clustered, dispersed, or random and provided a z-score and a P-value to aid the determination of whether or not to reject the null hypotheses that would indicate that the pattern was randomly distributed (*i.e.* spatially uncorrelated).

Results

The 1832 herald and main waves

Cholera first appeared in England during the autumn of 1831. By February 1832 it was epidemic in London. The 1832 LBoM recorded 3,200 cholera deaths. Of these, 582 occurred during the herald wave (between 21 February and 8 May), and then following a brief respite another 2,618 died during the main wave (between 15 May and 11 December). As a point of reference, the 1831 census reported that the population of parishes within the area of the LBoM at the time was just over one million (Census of Great Britain, 1831). Figures 2 and 3 show the geographic distribution of cumulative deaths during the herald and main waves at the ends of their respective exponential growth periods and their final ends.

A Spearman rank correlation compared the cumulative normalized cholera death count in each of the 55 parishes at the end of the herald's exponential growth period with that of each parish's normalized cumulative death count at the end of the main. The correlation showed a statistically significant relationship: r_s=0.39148, ρ=0.005. A second Spearman assessment comparing the parish cumulative normalized deaths at the ends of both waves similarly showed a statistically significant relationship: r_s=0.548624, ρ=0.001.

The parish normalized mortality provided the weighting factor in the determination of the weighted mean centre-point for the four epidemic segments. Figure 4 shows the centre-points on a large scale (small area) map. For comparison purposes the figure also shows the weighted centre of a combined herald/main wave, the geographic centre-point of all the parishes in the LBoM, and the weighted centre of the Sanitation Commission data. Although these data do not represent the same geographic area as the parish data nor the same demographics, they nevertheless provide a qualitative comparison of the two data sources. Also noteworthy is that all seven of the displayed centres are within 653 m of the centre of London Bridge. Spatial autocorrelation coefficients were determined for the four epidemic segments, and the resultant Moran's I for each (Table 1) shows a statistically significant clustered pattern.
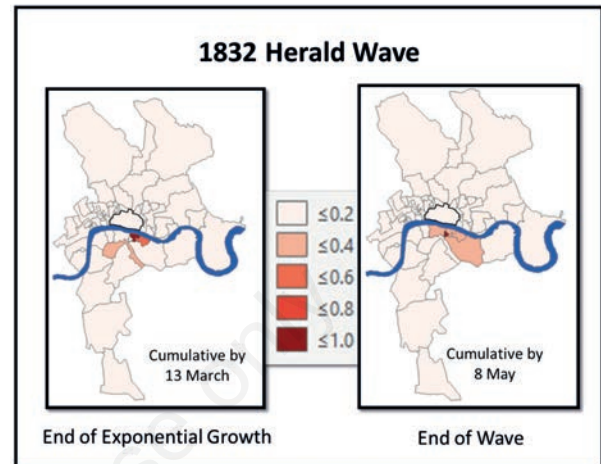

Figure 2. Geographic distribution of the number of cumulative deaths at the two essential herald wave points in 1832. The data have been normalized and then scaled into five bins of equal size as a percentage (×100) of the maximum observed Parish value during the wave segment.
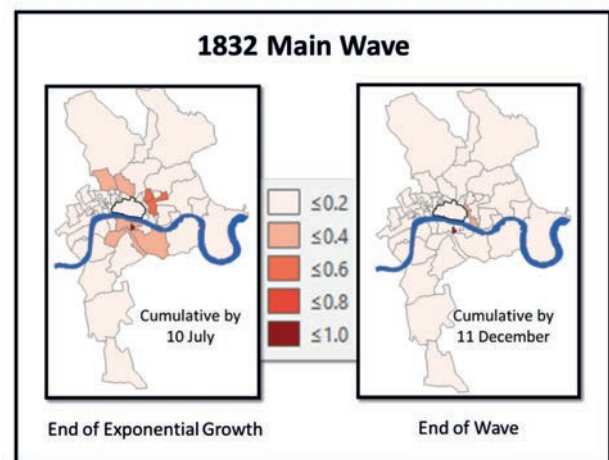

Figure 3. Geographic distribution of the number of cumulative deaths at the two essential main wave points in 1832. The data have been normalized and then scaled into five bins of equal size as a percentage (×100) of the maximum observed Parish value during the wave segment.

Table 1. Moran indices for 1832 herald and main wave segments.

1832 wave segments	Moran's I	Z	P
Herald - End of exponential growth	0.069069 clustered	2.587166	0.009677
Main - End of exponential growth	0.088888 clustered	2.491419	0.012723
Herald - End of wave	0.14002 clustered	4.656182	0.000003
Main - End of wave	0.035764 clustered	1.742268	0.081461

The 1848 herald and 1849 main waves

After 16 years of only unremarkable seasonal summertime cholera outbreaks following the 1832 epidemic, a non-seasonal outbreak occurred during the fall of 1848, peaking in January 1849. The herald wave continued to April of 1849 and left 1169 dead. The subsequent main wave began in May of 1849 and continued to December. The total cholera mortality during this main wave was a staggering 13,590. London's population had reached nearly two million according to the preceding census of 1841 (RGWR, 1848). Figures 5 and 6 show the geographic distribution of the normalized and scaled cumulative number of deaths during the herald and main waves at the ends of their respective exponential growth periods and at their wave's ends.

Wandsworth Registration District is London's most southwestern one. Its singularly high scaling shown at the end of the herald wave in Figure 5 resulted from a specific tragic situation. Mr. Drouet's Establishment for Pauper Children, located in Wandsworth, housed almost 1,400 children, and experienced a cholera outbreak in January 1849 that ultimately led to 180 deaths. This represents 15% of the herald wave's total London cholera mortality and obviously skewed the geographic data. Charles Dickens attributed the cause of the calamity at Drouet's establishment to the fact that it *'...was brutally conducted, vilely kept, preposterously inspected, dishonestly defended, a disgrace to a Christian community, and a stain upon a civilized land.'* (Dickens, 1849).

A Spearman rank correlation test was used to compare the normalized cumulative cholera death count in each of the 36 registration districts at the end of the herald's exponential growth period with that of each district's normalized cumulative death count at the end of the main's. The correlation showed a statistically significant relationship: $r_s = 0.366023$, $\rho = 0.025$. A second Spearman assessment comparing district normalized cumulative number of
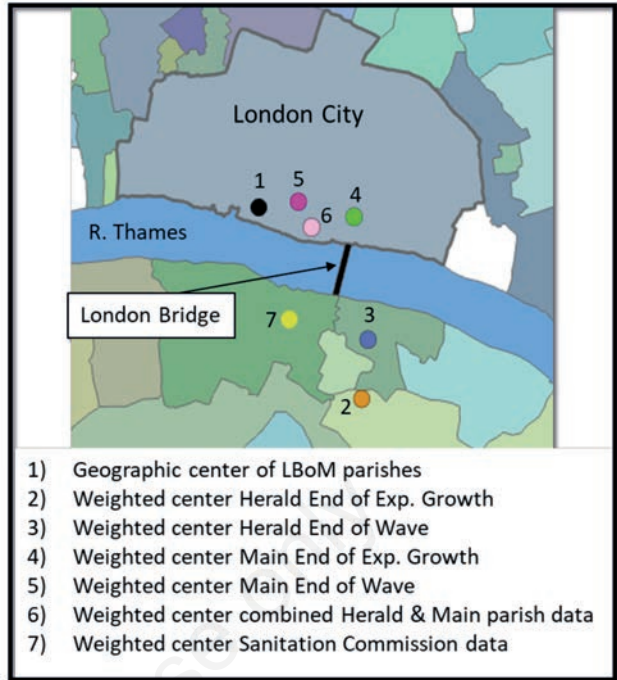

Figure 4. Relative locations of the 1832 wave segment centre-points. Geographic centre of the parishes in the Bills of Mortality in relation to the weighted-centre points of the herald and main waves' cumulative deaths at the end of their exponential growth periods and at their final ends. Also shown are the weighted centre points of the combined herald and main parish cumulative death counts and the weighted centre point of the cumulative deaths according to the Sanitation Commission.
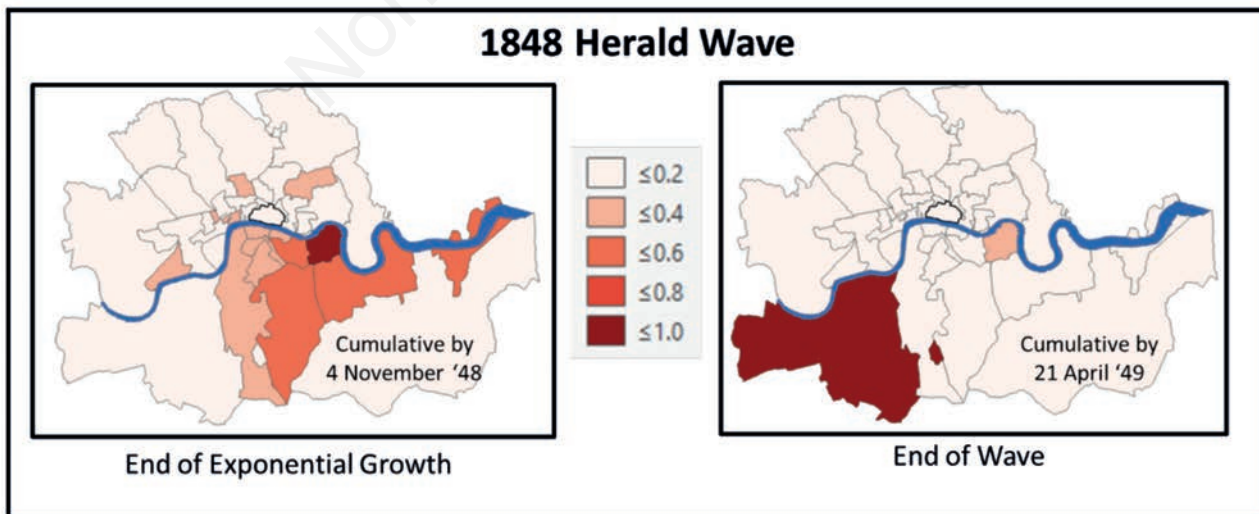

Figure 5. Geographic distribution of the number of cumulative deaths at the two essential herald wave points in 1848. The data have been normalized and then scaled into five bins of equal size as a percentage (×100) of the maximum observed Registration District value during the wave segment.

deaths at the ends of both waves also showed a statistically significant relationship: $r_s=0.432947$, $\rho=0.005$.

The normalized registration district mortality provided the weighting factor in the determination of a weighted mean centre point for the four epidemic segments. Figure 7 shows the centre-points on a large-scale map. It also shows the geographic centre of all the 36 Registration Districts. All four weighted wave segment centres are south of the Thames with the end of the herald wave centre separated from the other three clearly showing the effect of the outbreak at Drouet's establishment in Wandsworth. The weighted centres for the two end-of-exponential growth segments are both in the same Registration District. Spatial autocorrelation coefficients were determined for the four epidemic segments. The resulting Moran's *I* calculations show that three of the four segments have a statistically significant clustered pattern (Table 2). The scandalous situation at Drouet's school emerged in the '*random*' pattern calculated for the end of the herald wave.

The 1853 herald and 1854 main waves

It was only four years until the next non-summer outbreak occurred. This herald peaked in November of 1853 and lasted until the end of January 1854, taking 868 lives. The subsequent main began in July of 1854 and continued through the end of November. The total cholera mortality during this main wave was 10,668. The 1851 census showed the population in London's 36 registration districts at that time had reached almost 2.5 million (RGWR,

1853). The geographic distribution of normalized and scaled cumulative deaths during the herald and main waves at the ends of their respective exponential growth periods and at their final ends can be seen in Figures 8 and 9. The latter Figure also contains an illustration of a unique situation in St. James Westminster, a registration district north of the Thames, as it had a scale factor greater than 0.8, which is the only registration district (or parish) in any of the three epidemics having such a high value north of the Thames. It was this increased mortality that inspired the famous Dr John Snow to connect a presumed tainted water supply to the spread of cholera and to convince the Board of Guardians of St. James to remove the handle from its Broad Street water pump in September 1849 (Colwell, 1996). According to Colwell, this was the '*first instance on record of the implementation of an appropriate measure*' to stop the spread of cholera in a specific geographical district.

A Spearman rank correlation test was used to compare the normalized cumulative cholera death count in each of the 36 registration districts at the end of the herald's exponential growth period with that of each district's normalized cumulative death count at the end of the main wave's exponential growth period. The correlation showed a statistically significant relationship: $r_s=0.363028$, $\rho=0.025$. A second Spearman assessment comparing each district's normalized cumulative number of deaths at the ends of both the herald and main waves similarly showed a statistically significant relationship: $r_s=0.747748$, $\rho=0.005$.

The normalized mortality at the registration district level pro-

Table 2. Moran indices for 1848/49 herald and main wave segments.

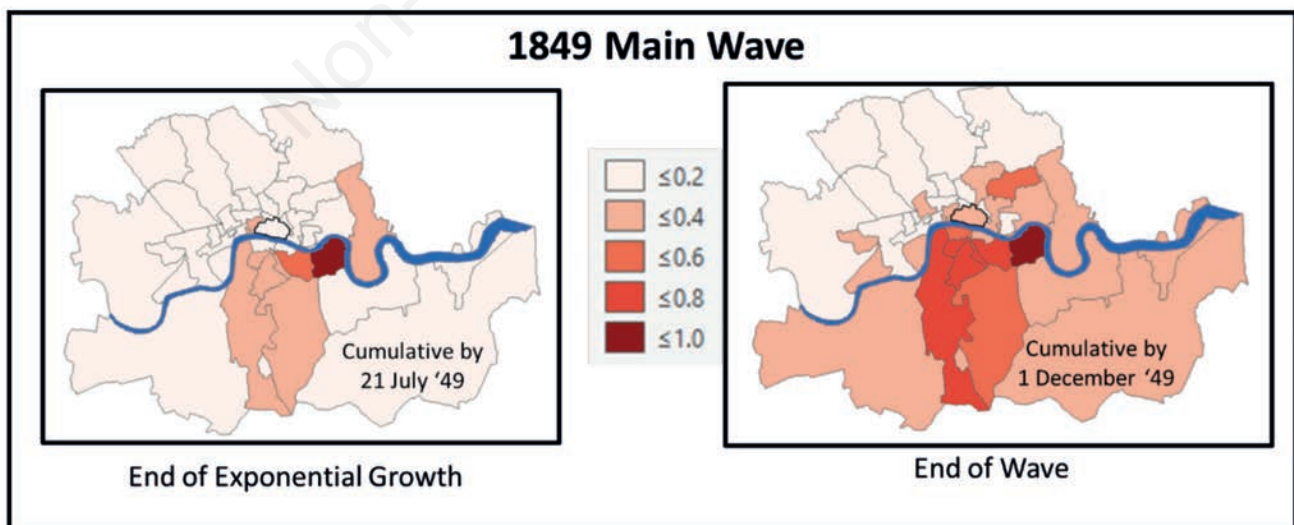| 1848/49 wave segments	Moran's *I*	Z	P
Herald - End of exponential growth	0.099421 clustered	2.740837	0.006128
Main - End of exponential growth	0.274197 clustered	7.517956	0
Herald - End of wave	−0.010236 random	−0.181621	0.855880
Main - End of wave	0.406644 clustered	10.460856	0

Figure 6. Geographic distribution of the number of cumulative deaths at the two essential main wave points in 1849. The data have been normalized and then scaled into five bins of equal size as a percentage (×100) of the maximum observed Registration District value during the wave segment.

vided the weighting factor in the determination of the weighted mean centre-point for the four epidemic segments (Figure 10).

The weighted centres of all four segments are south of the Thames. The centres of both end-of-exponential growth segments were found in adjacent registration districts, and so were also both the end-of-wave segment centres. Spatial autocorrelation coefficients were determined for the four epidemic segments. The Moran's *I* calculations (Table 3) showed that all four segments had a statistically significant clustered pattern.

Discussion

The 1832 epidemic

The statistically significant Spearman's rank correlation results are particularly compelling in demonstrating the similarity between this epidemic's herald and main waves. The end of exponential growth period comparison yielded a 99.5% confidence level (CL) that the correlation did not occur by chance. The correlation of the two waves at their ends resulted in an astonishing 99.9% CL. The Moran's *I* calculations show that the four wave segments were statistically clustered, further supporting the argument that the two waves were spatially similar.

The cholera mortality distributions shown in Figures 2, 3, and 4, however, do not provide such as striking evidence of a similarity between the herald and the main waves in each case. The herald seems to have been more concentrated south of the Thames, while the main was somewhat more concentrated to the north of the river.
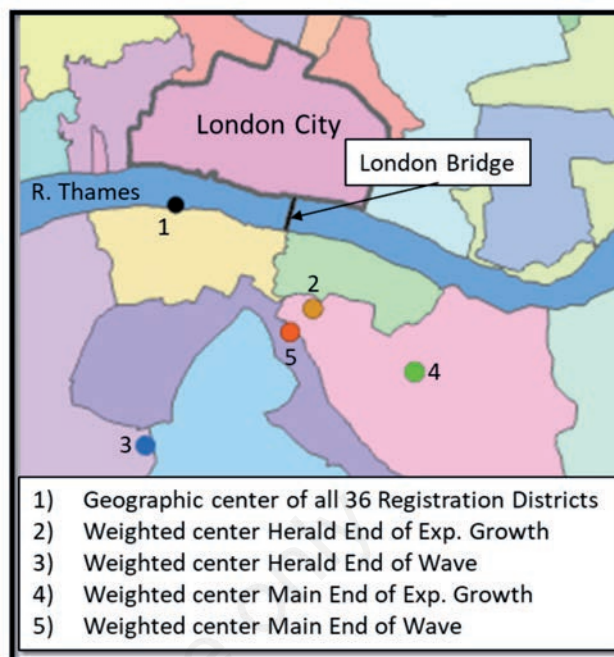

Figure 7. Relative locations of the 1848/49 wave segment centre-points. The geographic centre of the 36 Registration Districts in relation to the weighted-centre points of the herald and main waves' cumulative number of deaths at the end of their exponential growth periods and at the final ends.
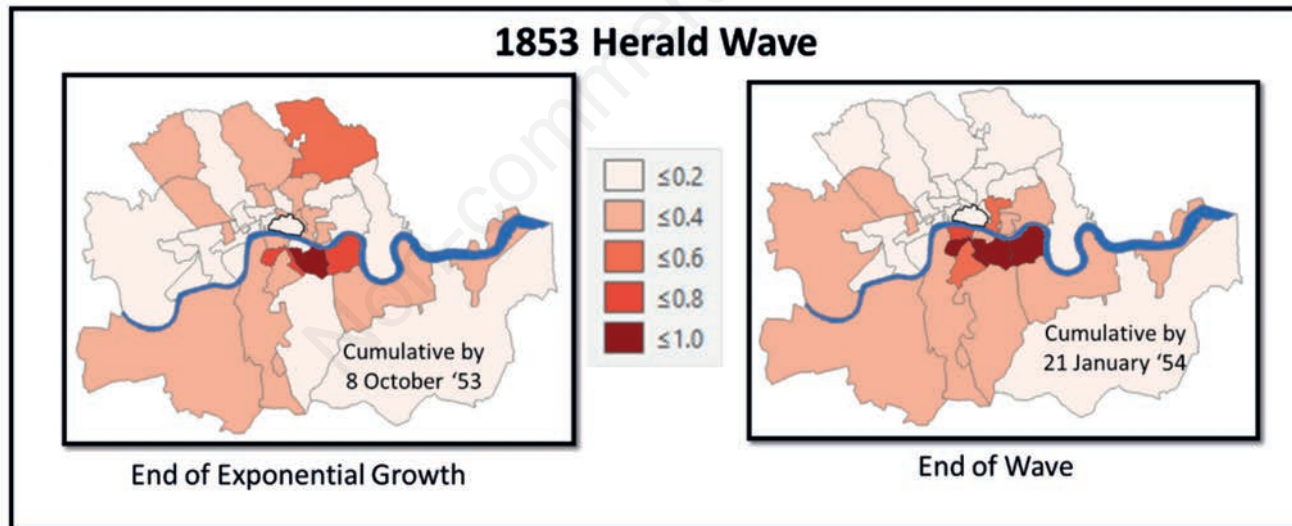

Figure 8. Geographic distribution of the number of cumulative deaths at the two essential herald wave points in 1853. The data have been normalized and then scaled into five bins of equal size as a percentage (×100) of the maximum observed Registration District value during the wave segment.

Table 3. Moran indices for 1853/54 herald and main wave segments.

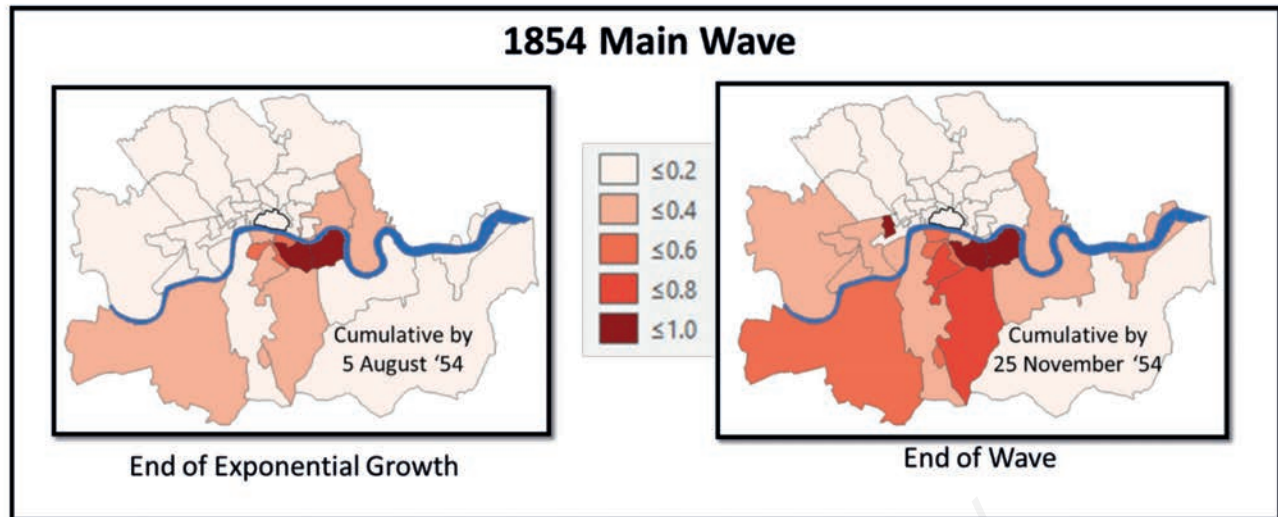| 1853/54 wave segments	Moran's *I*	Z	P
Herald - End of exponential growth	0.157181 clustered	4.191746	0.000028
Main - End of exponential growth	0.314858 clustered	8.299267	0
Herald - End of wave	0.313656 clustered	8.205458	0
Main - End of wave	0.31296 clustered	8.188801	0

Figure 9. Geographic distribution of the number of cumulative deaths at two essential main wave points in 1854. The data have been normalized and then scaled into five bins of equal size as a percentage (×100) of the maximum observed Registration District value during the wave segment.

Nevertheless, the weighted centre-points of all four segments were within 653 m of the centre of London Bridge and all parishes with a scaling factor of >0.8 were south of the river and none to the north. Deficiencies in the LBoM data are a possible reason that the depicted geographic distributions and segment weighted centre-points did not provide stronger evidence of similarity. The difference between the reported number of deaths '*within the walls*' in the LBoM and the number reported by the Sanitation Commission has already been cited. The difference manifested itself not only in the total number of epidemic deaths but, just as importantly, in their time of occurrence. Using the mortality '*within the walls*' as an example again, none of the 152 total deaths in the LBoM for the 1832 epidemic were recorded to have occurred during the herald; all were reported to have occurred during the main.

The 1848-49 epidemic

The statistically significant Spearman's rank correlation results are persuasive in demonstrating the similarity between this epidemic's herald and main waves. The end of the exponential growth period comparison yielded a 97.5% CL that the correlation did not occur by chance. The correlation of the two waves at their ends was equally persuasive at the 99.5% CL. The Moran's *I* calculations showed the effect of the singular cholera outbreak at Drouet's establishment in Wandsworth; while the herald end of exponential growth segment, the main end of exponential growth segment and the main wave end segment, all demonstrated significantly clustered results with significant z scores; it was only the herald's end-of-wave segment calculation that returned a '*random*' classification. The cholera mortality geographic distributions shown in Figures 5, 6 and 7 tend to support the argument that the herald and main waves of this epidemic were similar. The weighted centre-points of both end-of-exponential growth periods of both waves were in the same registration district south of the Thames. The weighted centre of the herald's end, despite the outbreak in Wandsworth, was only 1,400 m from the main's. Each registration district having a scaled factor >0.8 was situated south of the river Thames.

1853-54 epidemic

Again, the Spearman rank correlations comparing normalized cholera mortality of the herald with that of the main at the end of their exponential growth periods and their ends showed strong statistical correlations. The 97.5% CL for the end of exponential growth period comparison shows that the result was not due to chance, and 99.5% CL for the end of wave comparison was even
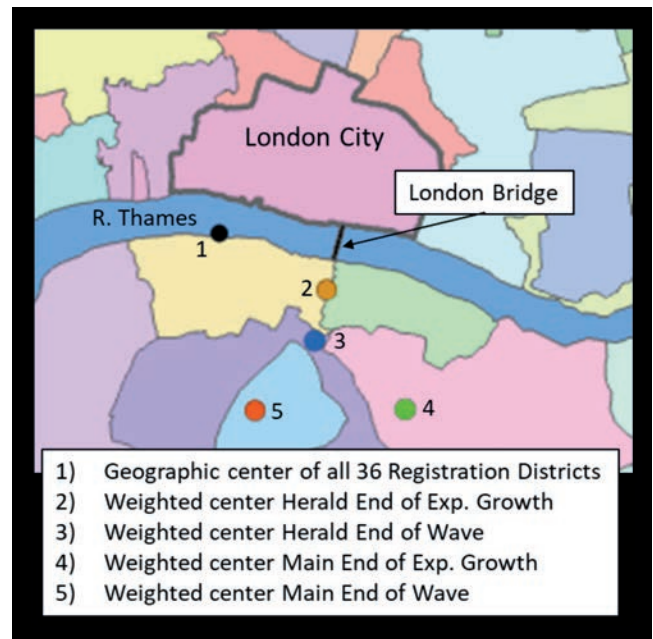

Figure 10. Relative locations of the 1853/54 wave segment centre-points. Geographic centre of the 36 registration districts in relation to the weighted-centre points of the number of cumulative deaths during the herald and main waves at the end of their exponential growth periods and at the final ends.

higher. Moran's *I* calculations for all four wave segments showed significantly clustered results. The minimum value of the four calculated results was 4.19 with the other three resulting in >8.0. The geographic distributions shown in Figures 8, 9, and 10 support the argument that the herald and main waves were spatially similar. The centre-points for the four segments were south of the Thames, with centre-points of both ends of the exponential growth periods in adjacent registration districts as were the ends of these waves. Also noteworthy is that all registration districts having a scaled factor >0.8 were south of the river except, of course, for the one containing the Broad Street pump. It was later determined that a cesspit adjacent to the pump, contaminated by washing the diapers of an infant victim, caused the situation (Whitehead, 1865).

Further investigation is required to determine if the observed herald cholera waves represent infections by a cholera strain different from that of the main waves. Advances in paleogenetics (Higuachi *et al.*, 1984; Pääbo, 2004) are promising, and recent research has demonstrated retrieval of *'…whole ancient pathogen genomes from archaeological specimens'* (Spyrou *et al.*, 2019) that can identify historically extant specific strains of such pathogens. Epidemiologists recognize the importance of geographical models to predict cholera epidemics (Colwell, 1996; Constantin de Magny *et al.*, 2008). An understanding of wave spatial characteristics would be a factor in differentiating between wave cholera strains and contribute to controlling future outbreaks.

Conclusions

Three of the four major cholera epidemics that swept London during the nineteenth century consisted of a non-summer outbreak - a herald wave - followed by a much more severe summer outbreak-a main wave. This study has explored the spatial characteristics of both these waves of these three epidemics with a view to determine if spatial differences existed that might possibly be related to the significant differences in the waves' severity. Notably, no spatial differences were identified between any of the herald waves and their subsequent main waves.

References

Burton N, Southall HR, 2004. GIS of the Ancient Parishes of England and Wales, 1500-1850. UK Data Service, SN 4828.

Census of Great Britain, 1831. Comparative account of the population of Great Britain in the years 1801, 1811, 1821 and 1831.

Dickens C, 1849. The Paradise at Tooting. London Examiner, 1/20/1849. Available from: https://www.thecircumlocutionoffice.com/theparadiseattooting/

Colwell R, 1996. Global climate and infectious disease: the cholera paradigm. Science 274:2025-31.

Constantin de Magny G, Murtugudde R, Sapiano M, Nizam A, Brown C, Busalacchi A, Yunus M, Nair G, Gil A, Lanata C, Calkins J, Manna B, Rajendran K, Bhattacharya M, Huq A, Sack R, Colwell R, 2008. Environmental signatures associated with cholera epidemics. PNAS 105:17676-81.

Higuchi R, Bowman B, Freiberger M, Ryder O, Wilson A, 1984. DNA sequences from the quagga, an extinct member of the horse family. Nature 312:282-4.

Ma J, Dushoff J, Bolker BM, Earn DJD, 2014. Estimating initial epidemic growth rates. Bull Math Biol 76:245-60.

Maplestone P, 2011. A lecture to the London & Middlesex Archaeological Society. Available from: https://www.london-parishclerks.com/History/The-Lectures.

Moran PAP, 1950. Notes on continuous stochastic phenomena. Biometrika 37:17-23.

Metropolitan Sanitary Commission, 1848. The Metropolitan Sanitary Commission First Report of the Commissioners. William Clowes & Sons, 428 pp.

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M, 2004. Genetic analyses from ancient DNA. Annu Rev Genet. 38:645-79.

Satchell AEM, Kitson PK, Newton GH, Shaw-Taylor L, Wrigley EA, 2018. 1851 England and Wales census parishes, townships, and places. UK Data Archive SN 852232. Available from: https://reshare.ukdataservice.ac.uk/852232/

Spyrou M, Bos K, Herbig A, Krause J, 2019. Ancient pathogen genomics as an emerging tool for infectious disease research. Genetics 20:323-40.

Tien JH, Poinar HN, Fisman DN, Earn DJD, 2011. Herald waves of cholera in nineteenth century London. J R Soc Interface 8:756-60.

Whitehead H, 1865. The broad street pump: an episode in the cholera epidemic of 1854. Macmillan's Magazine 13:113-22.

Worshipful Company of Parish Clerks, 1832. Weekly Bills of Mortality, London. Note: Weekly Bills of Mortality are bound together in annual volumes in London's Guildhall Library.