THE EUROPEAN SOCIETY OF MEDICINE
Medical Research Archives, Volume 12 Issue 9

RESEARCH ARTICLE

Application of tests for contagion in point processes to measles, Chlamydia, Lyme disease, and suicide

Ian McGovern[1], Frederic Schoenberg[1]

[1]UCLA Department of Statistics, 8125 Math-Science Building, Los Angeles, CA 90095-1554.

OPEN ACCESS

PUBLISHED
30 September 2024

CITATION
McGovern, I., Schoenberg, F., 2024. Application of tests for contagion in point processes to measles, Chlamydia, Lyme disease, and suicide. Medical Research Archives, [online] 12(9). https://doi.org/10.18103/mra.v12i9.5916

DOI
https://doi.org/10.18103/mra.v12i9.5916

ISSN
2375-1924

ABSTRACT

A recently proposed likelihood-ratio test for identifying causal triggering in point process data is applied to a variety of case counts of diseases of varying infectiousness. The test, suggested by McGovern et al. (2025), involves comparing the likelihood of a fitted Hawkes model to that of a fitted Poisson cluster model, and was shown using simulations to be powerful at discriminating between a process with causal triggering and a process where the clustering is merely due to spatial-temporal inhomogeneity. Here, the test is applied to data on measles, Chlamydia, and Lyme disease in the United States, to see if the test can discern between diseases that are highly contagious, moderately contagious, and not directly contagious from human to human. Measles is a highly contagious disease that spreads rapidly through populations, so it can potentially be modeled accurately using a Hawkes model[12]. Chlamydia is a sexually transmitted disease that is not as highly contagious as measles since the level of contact needed for exposure is much higher than for measles[2]. Lyme disease is non-contagious from human to human but cases tend to be highly clustered, as the disease is primarily spread through ticks, and this exposure is much more likely to happen during warmer weather[16]. Further, the test is applied to data on adolescent suicides in the United States, in order to investigate the hypothesis that such suicides are an epidemic spread by social contagion. The results show that the test is able to measure the degree of contagion of a disease, and the results suggest that there is indeed a small but statistically significant element of contagion to youth suicides.

Keywords: Epidemic disease, Hawkes process, Infectious disease, Inhomogeneous Poisson process, Poisson cluster process, test for triggering.

1. Introduction

A classical problem in the analysis of clustered point processes is the discrimination between triggering, where one point causes other points to be more likely to occur nearby, and inhomogeneity, where the aggregation of points occurs because certain spatialtemporal locations are simply more likely to encounter points, perhaps due to certain features of the spatial-temporal domain, or to spatial-temporal variation in certain covariates. This problem was discussed in Diggle(2014)[8], who suggested repeated observation of the point process as the primary way to distinguish these two phenomena. Unfortunately, in the case of most applications to infectious disease, repeated observation is impossible. For instance, while suicides are often clustered spatial-temporally and have been modeled via contagion models[6], it is possible that this spatial-temporal clustering might be explained as resulting from the large variation in certain covariates such as poverty, mental health issues, and gun ownership, all of which are correlated with suicide rates.

Recently, McGovern et al. (2025) proposed a likelihood-ratio-based test for contagion in spatialtemporal point process data, and used simulations to show the effectiveness and quantify its power in a variety of scenarios. Some questions remain, however.

- How well does the test work in practice, with actual epidemic diseases?
- Can the proposed test be used accurately to distinguish an infectious disease from a non-infectious disease?
- Can the test accurately quantify the degree of contagion for a given disease?
- Does the result of the test suggest that suicide is indeed an epidemic with significant contagion, as has been posited by various authors?[6] The present paper aims to investigate these questions.

Here, we apply the likelihood-ratio test proposed by McGovern et al. (2025) to several very different point process datasets in an attempt formally to assess the presence of significant contagion. Specifically, we apply the likelihood-ratio test not only to suicide data, where the existence of contagion is uncertain, but also to a disease well-known to be very highly contagious in its spread from human to human (measles), a disease known to be contagious but less so than measles (Chlamydia), and a disease well-known not to be directly contagious from human to human (Lyme disease).

Social contagion is a broad theory that states that behaviors or illnesses that are not physically contagious can be spread through social networks. The exact definition or mechanism of social contagion can vary widely among different fields and researchers, but in general the generally causal nature of this phenomenon is stressed. This paper explores testing the assumption that clustered behavior is the direct result of spread through social networks. Specifically, social contagion is often used as an explanation for suicidal[53] or non-suicidal self-harm behavior[11] in adolescents. Studies that have challenged this narrative typically rely upon attempting to isolate a "social contagion" factor amongst other variables that could lead to suicide[1], but such a method relies upon the assumption that no significant unobservable variables are confounding the results, which is not necessarily a reasonable assumption.

With regard to behaviors such as suicide, if there is an element of social contagion, then the spread could perhaps be modeled well with the same models that are used on epidemiological data for contagious diseases. A common model that is used on disease data within the point process paradigm is a Hawkes process[10], a self-exciting point process that lends itself well to modeling contagion. A Hawkes process consists of two elements - a background rate at which points are entering the spatial-temporal region at random, and a triggering element allowing points to increase the probability of additional points occurring in the future. The conditional intensity of such a temporal Hawkes model is of the form

$$\lambda(t|\mathcal{H}_t) = \mu + \kappa \sum_{i:t_i<t} g(t-t_i).$$

The background element is represented in the μ parameter, and the parameter κ, called the *productivity*, represents the expected number of points that every point will cause through triggering. The function $g(\Delta_t)$ is called the triggering density. For instance, if data are weekly aggregates and the geometric distribution is used for g, then $g(k) = p(1-p)^k$, for $k = 0,1,2,...$ weeks.

Within a contagious disease model, μ would typically represent the background rate of immigration of the disease into the population of interest, and the triggering element would represent spread between individuals. The parameter κ represents the speed at which the disease is spreading on average-a higher κ value corresponds to a more highly contagious disease.

Simply fitting a Hawkes model to adolescent suicide data would not necessarily definitively determine whether or not there is evidence of contagion within the data. Hawkes models can fit well to clustered data, regardless of the true clustering mechanism. That is, even if the aggregation of points in the point process is purely the result of inhomogeneity in explanatory variables, a Hawkes model representing triggering of points might nevertheless offer satisfactory fit. For this reason, more specialized methods are needed to distinguish between contagious clustering and clustering that is the result of inhomogeneity or non-causal clustering.

Diseases with known contagion methods can be compared to suicide data in order to evaluate and compare the fit of different clustering models. Three diseases that can be compared to suicide data in adolescents in order to study the social contagion theory are measles, Chlamydia, and Lyme disease. Measles is a highly contagious disease that spreads rapidly through populations, so it can potentially be modeled accurately using a Hawkes model[12]. Chlamydia is a sexually transmitted disease that is not as highly contagious as measles since the level of contact needed for exposure is much higher than for measles[2]. Lyme disease is non-contagious from human to human but cases tend to be highly clustered, as the disease is primarily spread through ticks, and this exposure is much more likely to happen during warmer weather[16].

2. Materials and Methods

Several tests for distinguishing between triggering and inhomogeneity were summarized in McGovern et al. (2025) and are briefly reviewed here. In order to assess whether or not a causal (contagious) clustering model such as a Hawkes model fits significantly better to data than a non-causal clustering model, a hypothesis test using the information gain statistic is used. The information gain statistic measures the predictive properties of a point process model[7] and is given by

$$\hat{I} = \frac{1}{n}(\log(L_1) - \log(L_0)).$$

where n is the number of points in the point process , and L_1 and L_0 represent the likelihood of the point process under the alternative and null model, respectively[9].

While the null hypothesis is that N is an inhomogeneous Poisson process and the alternative is that N is a Hawkes process, the test statistic proposed in McGovern et al. (2025) is the log-likelihood difference between a fitted Poisson cluster model[13] and a fitted Hawkes model. A Poisson cluster model is a noncausal clustering model consisting of a hidden layer of "parent" points that trigger "children" according to some triggering density, where each child is distributed randomly about its parent. In the purely temporal version with triggering density symmetric about zero, the children are equally likely to occur prior to their parent or after their parent. Thus, the Poisson cluster model can be considered non-causal, as points triggered backwards in time contradict the behavior of a contagious or causal clustering model. In McGovern et al. (2025) this statistic was shown to be powerful at discriminating a Hawkes process from an inhomogenous Poisson process. The idea is that, if the underlying mechanism is an inhomogeneous Poisson process,

then a Poisson cluster model would be expected to fit as well as a Hawkes process, whereas if the process truly exhibits causal triggering, then a Hawkes model should offer superior fit.

A temporal Poisson cluster process N can be viewed as an example of a Cox process, where the intensity is random, as it depends on the random collection M of hidden parent points, but given M, the process N is a Poisson process with intensity

$$\lambda_{NS}(t|M) = A \sum_{i:t_i \in M} h(t - t_i).$$

The parameter A represents the average number of points that each parent triggers, and $h(\Delta_t)$ represents the clustering density. Under this parameterization, this conditional intensity is not based on the history of the process N but instead is based on knowing the hidden parent process M. In addition, the clustering function $h(\Delta_t)$ is not limited to positive values unlike the triggering density in a Hawkes model.

In order to estimate the sampling distribution of the information gain statistic under the null, McGovern et al. (2025) propose a Monte Carlo method.

Specifically, in the context of the datasets analyzed in the present analysis, a Poisson cluster process is first fit to the data using a Gaussian triggering density. A Gaussian clustering algorithm is then performed using maximum likelihood estimation, and the parameters for the Poisson cluster process are estimated using the number of clusters (μ), the mean number of points per cluster (A), and the standard deviation of the clusters (σ). Poisson cluster processes with these same parameters are then simulated, and a Hawkes model is fit to each Poisson cluster simulation by maximum likelihood. The information gain statistic is then calculated, and the sampling distribution used is the collection of information gains between the Hawkes and Poisson cluster model log-likelihoods for all of the simulations. Here, 500 simulations were used for each dataset. The information gain statistic is then calculated for the original data in the same manner and compared to the simulated sampling distribution to determine if the null hypothesis can be rejected.

Since the datasets each consist of weekly case counts, for the Hawkes model, a geometric distribution was used for the triggering density, which has been used to model daily Covid-19 case counts in a similar situation.[4] In fitting the Hawkes model for each dataset, the parameters μ, κ and p are fit using maximum likelihood. The square root of the diagonal elements of the estimated Hessian of the log-likelihood was used for standard error estimates of the parameters in fitted Hawkes models, as suggested by Ogata[14].

2.1 DATA

Adolescent suicide statistics were collected from the CDC Wonder Provisional Mortality Statistics database which begins at 2018 and collects data through the most recent week. The provisional mortality statistics are based upon death certificates for United States residents, and any category with less than nine deaths is suppressed, as the CDC cannot guarantee the accuracy of low numbers of deaths. No suppressed results were included in this analysis.

The age range selected was from the five year age groups of 10-14 and 15-19. The method of death was limited to UCD- ICD-10 Codes of X60-X84, which are all intentional self-harm deaths. Specific method of self-harm death was not considered. The data selected was based upon the years that the CDC did not label as "provisional" in their reporting, which are the dates 2018-01-01 to 2021-31-12. The data was collected weekly as this was the smallest time interval is released, and there were no suppressed values in the output.

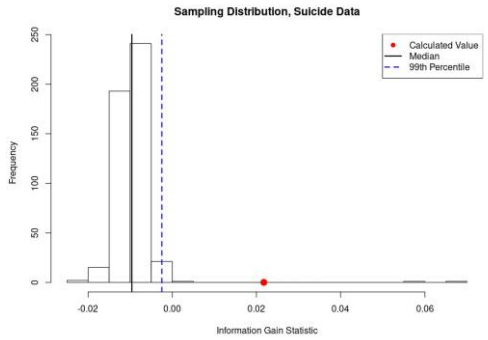
Measles data[19], Chlamydia data,[17] and Lyme data[18] for the United States were collected from Project Tycho, which compiles weekly case counts of various diseases[15], with some geographical information included.

Measles data consisted of weekly measles cases in Los Angeles County from 1928-01-01 to 1931-31-12 collected from Project Tycho. The length of the data collected was chosen to match the number of

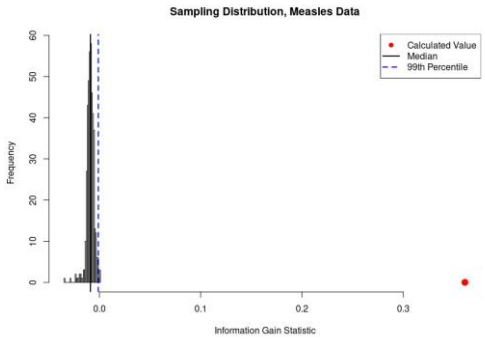
weeks available for the adolescent suicide data, to make sure the power of the tests were similar. The data was limited to Los Angeles as different counties in California reported at different frequencies, and Los Angeles county reported data consistently. The time period was selected as the first 4-year time period in which data was consistently reported weekly.

Lyme data compiled in Project Tycho consisted of weekly case counts of Lyme disease in California from 2008-01-01 to 2011-12-31.The four year period matches the length of the observed data for adolescent suicide deaths. The years 2008 to 2011 were selected for this analysis as this was the first four year time period to have consistent data reported weekly. The data were limited to the state of California, which had the most consistent reporting of weekly case counts, with no weeks missing during 2008-2011.

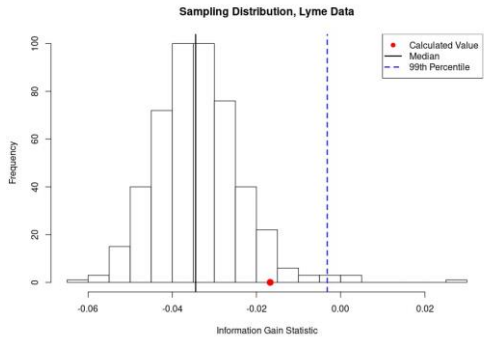
Weekly case counts of Chlamydia in California from 2008-01-01 to 2011-12-31 were compiled in Project Tycho. The time period and state was chosen to match the Lyme disease data, and the Chlamydia data also had the same consistent weekly case reporting in California during that time period.



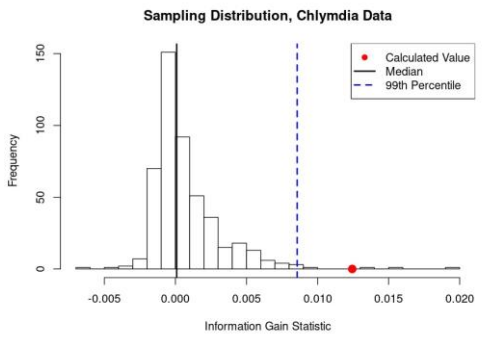
(a) Adolescent Suicides



(b) Measles



(c) Lyme



(d) Chlamydia

Figure 1: Log-likelihood test statistic and simulated null distribution for (a) adolescent suicide data in the United States from 2018-01-01 to 2021-31-12; (b) reported measles cases in Los Angeles County, from 192801-01 to 1931-31-12; (c) reported Lyme cases in California, from 2008-01-01 to 2011-31-12; (d) reported Chlamydia cases in California, from 2008-01-01 to 2011-31-12

3. Results

For measles, the estimated value of the information gain statistic is well above any of the values in the sampling distribution, as seen in Figure 2a. The p-value is essentially 0, so the null hypothesis is therefore rejected. This corresponds to the expected result, since measles is highly contagious. The estimated value of the κ parameter is 0.973 with a confidence interval of (0.954, 0.993), indicating significant clustering in the Hawkes process.

Table 1: Results of Tests for Clustering and Causal Clustering

	κ(99% CI)	p-value of Hypothesis Test
Measles	(0.954 0.993)	.000
Chlamydia	(.569, .591)	.006
Lyme Disease	(.429,.720)	.061
Adolescent Suicide	(-.012,.269)	.004

For the Chlamydia data, the p-value of the estimated information gain statistic is 0.006, so the null hypothesis is again rejected. The estimated κ parameter is 0.580 with a 99% confidence interval of (0.569,0.591).

For Lyme disease, the estimated information gain statistic has a p-value of 0.061, indicating that the null hypothesis is not rejected. The maximum likelihood estimate of κ is 0.575 with a 99% confidence interval of (0.429, 0.720).

The estimated information gain statistic applied to the youth suicide data results in a p-value of 0.004, indicating that the null hypothesis is rejected, though the corresponding κ estimate is just 0.128 with a 99% confidence interval of (−.012,.269).

As seen in Figure 3b, the Poisson cluster model fails to accurately account for the large outbreak of the disease that occurs during the observation window. As with measles, one sees in Figure 3d that the Poisson cluster model offers much poorer fit to the largest outbreak of Chlamydia, compared to the fitted Hawkes model. Figure 3c shows that the Hawkes and Poisson cluster models offer very similar fit. As shown in Figure 3a, the Hawkes model does appear to fit better perhaps due to the Poisson cluster model overfitting to the small level of clustering that is present. The fitted Hawkes model, by contrast, has only mild fluctuations in conditional intensity over the observation window. Even though it appears there is some level of clustering in the suicide data as seen in Figure 2a, this clustering is not as significant as the clustering of the other conditions examined here.

4. Discussion

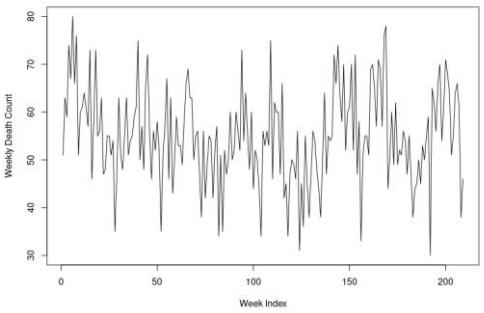
Of the four conditions considered here, measles, Chlamydia, Lyme disease, and suicide, the likelihood ratio tests for measles, Chlamydia, and suicide suggest that the Hawkes model fits significantly better than a Poisson cluster model. The only condition for which the likelihood ratio test did not reject the null hypothesis was Lyme disease. This suggests that the likelihood ratio test used is accurately measuring contagion, since measles is known to be the most highly contagious disease of the four studied here. In fact, the information gain statistic is very clearly higher than any of the values that were drawn from the sampling distribution, indicating that the Hawkes process does fit much better to the measles data as compared to a non-causal Poisson cluster model. The measles condition also has the highest κ estimate, which is consistent with the highly contagious nature of measles.

The log-likelihood ratio test statistic for Chlamydia is also highly significant, with an estimated pvalue of 0.006. The estimated κ parameter is also much lower than the corresponding estimate for measles, which is reasonable due to the nature of Chlamydial infection and its known degree of contagion. Despite the lower level of infection for Chlamydia, the Hawkes model nevertheless provides a significantly better fit than the Poisson cluster model.

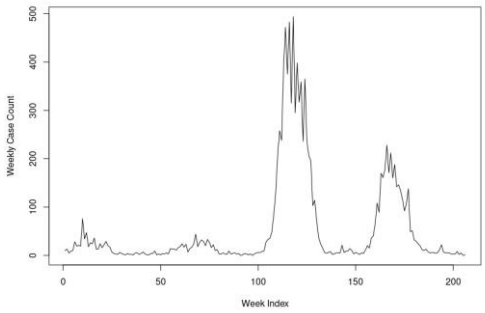
For the Lyme data, the likelihood ratio test statistic is not statistically significant, with an estimated p-value of 0.061. The Lyme data also has a κ parameter estimate that is statistically significantly different from 0, and is around the same value as the κ statistic from the Chlamydia data. However, the p-value and the much larger uncertainty

regarding the κ value both provide doubt that the clustering within the Lyme data is evidence of contagion. The Poisson cluster model fits well to the data on Lyme disease, a result which is consistent with the fact that Lyme disease exhibits seasonal patterns but is not directly infectious from human to human.

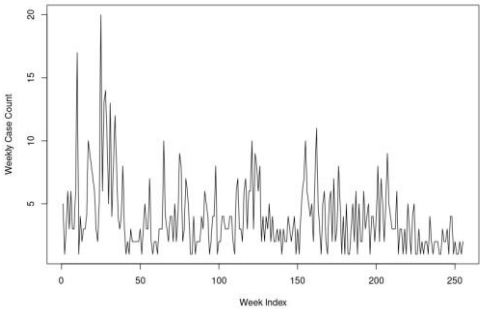
The one set of data where the disease mechanism is most uncertain is the adolescent suicide data. According to the social contagion paradigm, a Hawkes model should fit this data closely. In fact, the suicide data is well approximated by a Hawkes model, significantly better than by a Poisson cluster model.



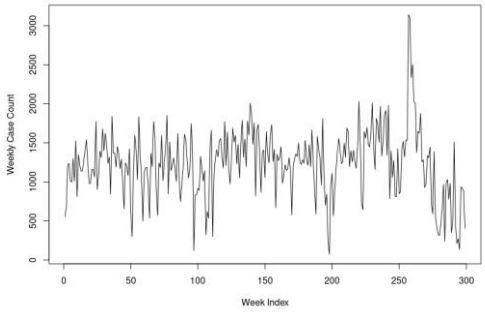
(a) Adolescent Suicide



(b) Measles



(c) Lyme



(d) Chlamydia

Figure 2: Weekly reported case counts for (a) adolescent suicide deaths in the United States, from 2018-01-01 to 2021-31-12; (b) measles cases in Los Angeles County, from 1928-01-01 to 1931-31-12; (c) Lyme disease cases in California, from 2008-01-01 to 2011-31-12; (d) Chlamydia cases in California, from 2008-01-01 to 2011-31-12.

However, the estimated κ parameter for the suicide data is very low and not statistically significant. This indicates that the best fitting Hawkes model has essentially just a constant background rate parameter μ and very little triggering.

For suicide data, while the likelihood ratio test is significant, the estimated level of contagion is very low. Thus the results here suggest that there is a small but statistically significant element of contagion to the youth suicide data. It may be that, while there are many far more important predictors and explanations for youth suicides, one minor contributor is the knowledge of and exposure to other youth suicides.

Conflict of Interest:

There are no conflicts of interest with regards to this paper's two authors, Ian McGovern and Frederic Schoenberg.

Funding Statement:

This material is based upon work supported by the National Science Foundation under grant number DMS-2124433.

Acknowledgements:

Thanks to Project Tycho and CDC Wonder for supplying the data.

References:

[1] M. Ali, D. Dwyer, and J. Rizzo. "The social contagion effect of suicidal behavior in adolescents: does it really exist?" In: *The journal of mental health policy and economics* 14.1 (Mar. 2011), pp. 3–12. issn: 1091-4358. url: http://europepmc.org/abstract/MED/21642746

[2] C.L. Althaus, K. Turner, Schmid B., Low N., J.C.M Heinje, and M. Kretzshmar. "Transmission of Chlamydia trachomatis through sexual partnerships: a comparison between three individual-based models and empirical data". In: *R. Soc. Interface* (2012).

[3] P.S. Bearman and J. Moody. "Suicide and Friendships Among American Adolescents". In: *American Journal of Public Health* 94.1 (2004). PMID: 14713704, pp. 89–95. doi: 10.2105/AJPH.94.1.89. eprint: https://doi.org/10.2105/AJPH.94.1.89. url: https://doi.org/10. 2105/AJPH.94.1.89.

[4] R. Browning, D. Sulem, K. Mengersen, V. Rivoirard, and J. Rousseau. "Simple discretetime self-exciting models can describe complex dynamic processes: A case study of COVID19". In: *PLOS ONE* 16.4 (Apr. 2021), pp. 1– 28. doi: 10.1371/journal.pone.0250015. url:https://doi.org/10.1371/journal.pone.0250015.

[5] Q. Cheng, H. Li, V. Silenzio, and E. Caine. "Suicide Contagion: A Systematic Review of Definitions and Research Utility". In: *PLOS ONE* 9.9 (Sept. 2014), pp. 1–9. doi: 10.1371/journal.pone.0108724. url:https://doi.org/10. 1371/journal.pone.0108724.

[6] Q. Cheng, H. Li, V. Silenzio, and E.D. Caine. "Suicide Contagion: A Systematic Review of Definitions and Research Utility". In: *PLOS ONE* 9.9 (Sept. 2014), pp. 1–9. doi: 10.1371/ journal.pone.0108724. url:https://doi.org/10. 1371/journal.pone.0108724.

[7] D. J. Daley and D. Vere-Jones. "Scoring probability forecasts for point processes: the entropy score and information gain". In: *Journal of Applied Probability* (2016).

[8] P. J. Diggle. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press, 2014.

[9] D. Harte and D. Vere-Jones. "The Entropy Score and its Uses in Earthquake Forecasting". In: *Pure Applied Geophysics* 162 (2005).

[10] A. Hawkes. "Spectra of some self-exciting and mutually exciting point processes". In: *Biometrika* 58.1 (Apr.1971), pp. 83–90. issn: 0006-3444. doi: 10.1093/biomet/58.1.83. eprint: https://academic.oup.com/biomet/article-pdf/58/1/83/602628/58-1-83.pdf. url: https://doi.org/10.1093/biomet/58.1.83.

[11] S. Jarvi, B. Jackson, B. Swenson, and H. Crawford. "The Impact of Social Contagion on NonSuicidal Self-Injury: A Review of the Literature". In: *Archives of Suicide Research* 17.1 (2013). PMID: 23387399, pp. 1–19. doi: 10. 1080/13811118 .2013 . 748404 . eprint: https : //doi.org/10.1080/13811118.2013.748404. url: https://doi.org/10.1080/13811118.2013. 748404.

[12] B.M. Laksono, R.D. De Vries, S. McQuaid, W. P. Duprex, and R.L. De Swart. "Measles Virus Host Invasion and Pathogenesis". In: *Viruses* 8.8 (2016). issn: 1999-4915. doi: 10. 3390/v8080210. url: https://www.mdpi.com/ 1999-4915/8/8/210.

[13] J. Neyman and E. Scott. "Statistical Approach to Problems of Cosmology". In: *Journal of the Royal Statistical Society* 20.1 (1958), pp. 1–43.

[14] Y. Ogata. "The asymptotic behaviour of maximum likelihood estimators for stationary point processes". In: *Ann Inst Stat Math* 30 (1978), pp. 243–261. url: https://doi.org/10.1007/ BF02480216.

[15] W. van Panhuis. "Contagious diseases in the United States from 1888 to the present." In: *The New England journal of medicine* 369.22 (2013). doi: 10.1056/NEJMms1215400.

[16] A. Roome, R. Spathis, L. Hill, J.M. Darcy, and Ralph M. Garruto. "Lyme Disease Transmission Risk: Seasonal Variation in the Built Environment". In: *Healthcare* 6.3 (2018). issn: 22279032. doi: 10.3390/healthcare6030084. url: https://www.mdpi.com/2227-9032/6/3/84.

[17] W. Van Panhuis, A. Cross, and D. Burke. "Counts of Chlamydial infection reported in UNITED STATES OF AMERICA: 19952016". In:

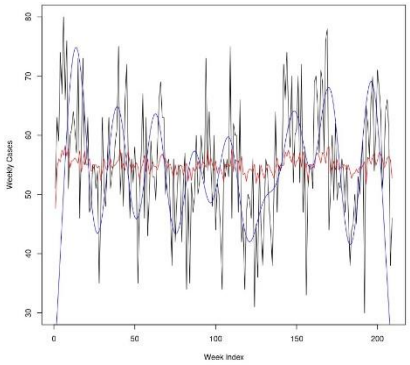
(Apr. 2018). doi: 10.25337/T7/ ptycho.v2.0/US.105629000. url:https://www. tycho.pitt.edu/dataset/US.105629000/.

[18] W. Van Panhuis, A. Cross, and D. Burke. "Counts of Lyme disease reported in UNITED STATES OF AMERICA: 1990-2016". In: (Apr. 2018). doi: 10.25337/T7/ptycho.v2.0/US. 23502006. url: https://www.tycho.pitt.edu/ dataset/US.23502006/.

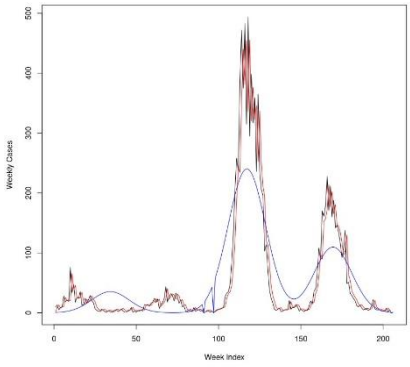
[19] W. Van Panhuis, A. Cross, and D. Burke. "Counts of Measles infection reported in UNITED STATES OF AMERICA: 1888-2002". In: (Apr. 2018). doi: 10.25337/T7/ ptycho.v2.0/US.14189004. url: https://www. tycho.pitt.edu/dataset/US.14189004/.

[20] McGovern I, Brantingham PJ, and Schoenberg F (2025). Testing for causal clustering in point processes. Journal of Applied Science and Innovation Studies, to appear.

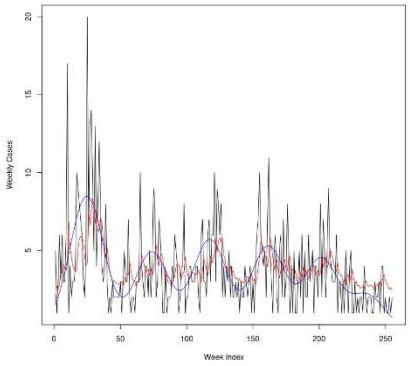
5. Appendix



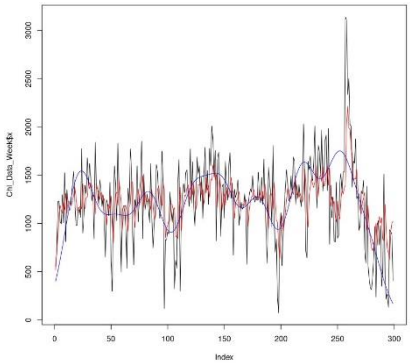
(a) Adolescent Suicide



(b) Measles

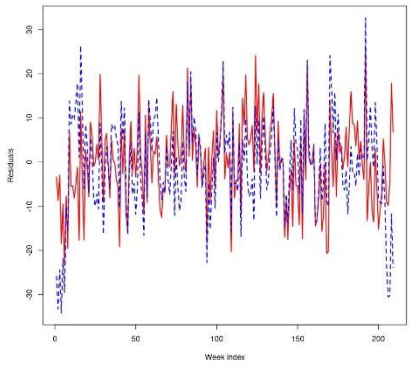


(c) Lyme

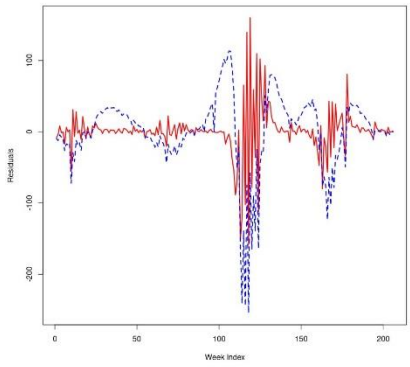


(d) Chlamydia

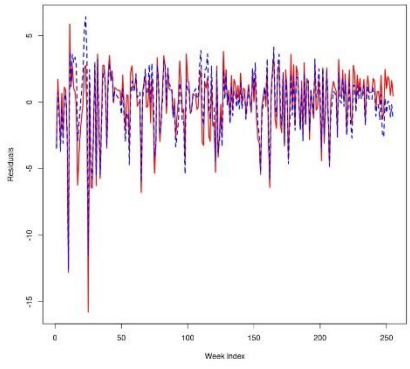
Figure 3: Projected case counts in fitted Hawkes (red) and Poisson cluster (blue) models for (a) Adolescent suicide data in the United States from 2018-01-01 to 2021-31-12; (b) measles cases in Los Angeles County, from 1928-01-01 to 1931-31-12; (c) Lyme cases in California, from 2008-01-01 to 2011-31-12; (d) Chlamydia Cases in California, from 2008-01-01 to 2011-31-12



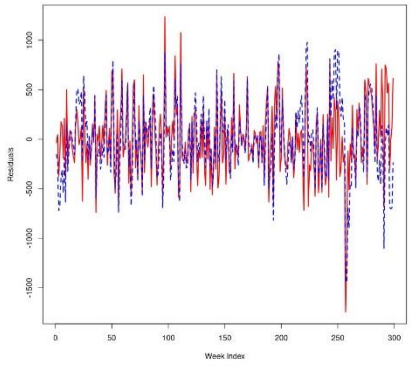
(a) Adolescent Suicides



(b) Measles

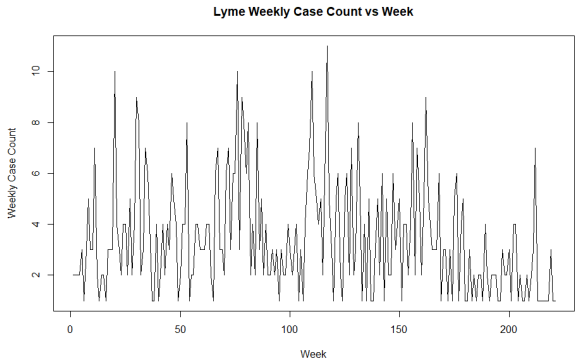


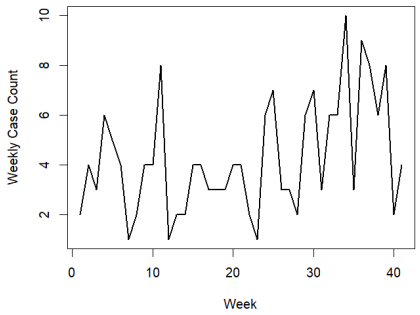
(c) Lyme

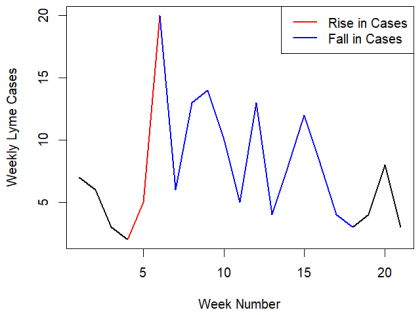


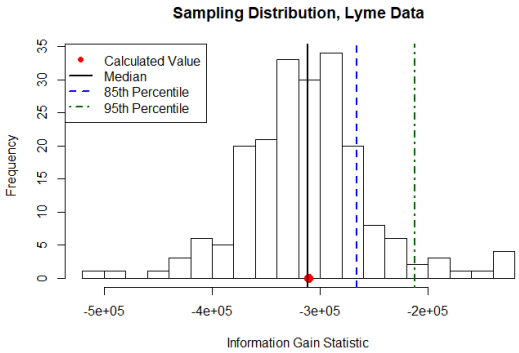
(d) Chlamydia

Figure 4: Comparison of Residuals of projected case counts in fitted Hawkes (red) and Poisson cluster (blue) models for (a) Adolescent suicide data in the United States from 2018-01-01 to 2021-31-12; (b) measles cases in Los Angeles County, from 1928-01-01 to 1931-31-12; (c) Lyme cases in California, from 2008-01-01 to 2011-31-12; (d) Chlamydia Cases in California, from 2008-01-01 to 2011-31-12


Lyme Weekly Case Count vs Week
Weekly Case Count
Week
0
50
100
150
200
2
4
6
8
10


Weekly Case Count
Week
0
10
20
30
40
2
4
6
8
10


Rise in Cases
Fall in Cases
Weekly Lyme Cases
Week Number
5
10
15
20


Sampling Distribution, Lyme Data
Calculated Value
Median
85th Percentile
95th Percentile
Frequency
Information Gain Statistic
-5e+05
-4e+05
-3e+05
-2e+05
0
5
10
15
20
25
30
35