Detecting changes in dispersion in COVID-19 case counts using a negative binomial model

JSM 2024

Rachael Aber

2024

Oregon State University

Table of Contents

Oregon State University

Table of Contents

Oregon State University

Why study variability?

- Highly variable case count time series suggest transmission heterogeneity(bc), demographic/environmental heterogeneity, or changes in R

Oregon State University

Why study variability?

- Highly variable case count time series suggest transmission heterogeneity(bc), demographic/environmental heterogeneity, or changes in R
- Metrics of variability are often an overlooked way to understand systems (How is variability related to different phases of an epidemic? [**graham˙measles˙2019**]

Oregon State University

Why study variability?

- Highly variable case count time series suggest transmission heterogeneity(bc), demographic/environmental heterogeneity, or changes in R
- Metrics of variability are often an overlooked way to understand systems (How is variability related to different phases of an epidemic? [**graham˙measles˙2019**]
- Adam et al. [**adam˙time-varying˙2022**] found that COVID-19 transmission heterogeneity decreased over time and was associated with interventions to slow spread

Oregon State University

Why study variability?

- Highly variable case count time series suggest transmission heterogeneity(bc), demographic/environmental heterogeneity, or changes in R
- Metrics of variability are often an overlooked way to understand systems (How is variability related to different phases of an epidemic? [**graham˙measles˙2019**]
- Adam et al. [**adam˙time-varying˙2022**] found that COVID-19 transmission heterogeneity decreased over time and was associated with interventions to slow spread
- Information about what phase/dynamic regime an epidemic is in, as well as potentially indicating the level of heterogeneity at finer spatial and temporal scales (in)

Oregon State University

Why study variability?

- Dispersion around a rolling mean (moving window) of a case count time series may be a useful metric because it forms part of a framework that models variance more flexibly than a Poisson model

Why study variability?

- Dispersion around a rolling mean (moving window) of a case count time series may be a useful metric because it forms part of a framework that models variance more flexibly than a Poisson model
- A 'mean crowding' parameter [**lloyd˙mean˙1967**] was proposed, which is the mean number per individual of other individuals in the same quadrat

Oregon State University

Why study variability?

- Dispersion around a rolling mean (moving window) of a case count time series may be a useful metric because it forms part of a framework that models variance more flexibly than a Poisson model
- A 'mean crowding' parameter [**lloyd˙mean˙1967**] was proposed, which is the mean number per individual of other individuals in the same quadrat
- Useful way to think about dispersion in case count time series, degree of dispersion is degree of clustering/crowding of cases (rel to crowdin parm) (from the perspective)

Oregon State University

Table of Contents

Oregon State University

Introduction to the method

- Investigated changes in dispersion over time (many processes may affect dispersion)

Oregon State University

Introduction to the method

- Investigated changes in dispersion over time (many processes may affect dispersion)
- Adjusted for population size using an offset in the model (directly model)

Oregon State University

Introduction to the method

- Investigated changes in dispersion over time (many processes may affect dispersion)
- Adjusted for population size using an offset in the model (directly model)
- Dispersion allowed to vary more slowly than the process mean

Oregon State University

Introduction to the method

- Investigated changes in dispersion over time (many processes may affect dispersion)
- Adjusted for population size using an offset in the model (directly model)
- Dispersion allowed to vary more slowly than the process mean
- Linear predictor includes a natural spline in time to account for autocorrelation in case counts (ns are)

$$log(E[Y_i]/n_i) = \beta_1 h_1(t_i) + \beta_2 h_2(t_i) + \beta_3 h_3(t_i) \quad (1)$$

$$log(E[Y_i]) - log(n_i) = \beta_1 h_1(t_i) + \beta_2 h_2(t_i) + \beta_3 h_3(t_i) \quad (2)$$

$$log(E[Y_i]) = \beta_1 (h_1(t_i) + \beta_2 h_2(t_i) + \beta_3 h_3(t_i) + log(n_i) \quad (3)$$

Oregon State University

Negative binomial model

- $$f_t(I) = \binom{I+\theta-1}{I} \frac{\mu}{\mu+\theta}^{I} \frac{\theta}{\mu+\theta}^{\theta} \quad (4)$$

Oregon State University

Negative binomial model

- $$f_t(I) = \binom{I+\theta-1}{I} \frac{\mu}{\mu+\theta}^{I} \frac{\theta}{\mu+\theta}^{\theta} \quad (4)$$

- $$E(I) = \mu \quad (5)$$

$$Var(I) = \mu + \frac{\mu^2}{\theta} \quad (6)$$

Oregon State University

Application to simulated data

- For validity/power simulations, we used both Gaussian and uniform epidemic curves with an attack rate of 0.1

Application to simulated data

- For validity/power simulations, we used both Gaussian and uniform epidemic curves with an attack rate of 0.1
- Epidemic curves over 60 timesteps each were produced, and a likelihood-ratio test (LRT) procedure was applied to each

Oregon State University

Application to simulated data

- For validity/power simulations, we used both Gaussian and uniform epidemic curves with an attack rate of 0.1
- Epidemic curves over 60 timesteps each were produced, and a likelihood-ratio test (LRT) procedure was applied to each
- Varying the effect size, location of the breakpoint, population size, and curve shape allowed us to test the validity and power of our approach

Oregon State University

Application to empirical data

- Mention diagnostic plots/dispersion statistics (take out?)

Oregon State University

Application to empirical data

- Mention diagnostic plots/dispersion statistics (take out?)
- We estimated μ_t and θ_t using an iterative reweighted least-squares (procedure implemented via the NBPSeq package[**NBPSeq**] and from Di et al.[**yanming˙nbp˙2011**] with a moving window approach (for each window)

Oregon State University

Application to empirical data

- Mention diagnostic plots/dispersion statistics (take out?)
- We estimated μ_t and θ_t using an iterative reweighted least-squares (procedure implemented via the NBPSeq package[**NBPSeq**] and from Di et al.[**yanming˙nbp˙2011**] with a moving window approach (for each window)
- We investigated large counties (largest three counties in each state), due to power constraints

Oregon State University

Table of Contents

Oregon State University

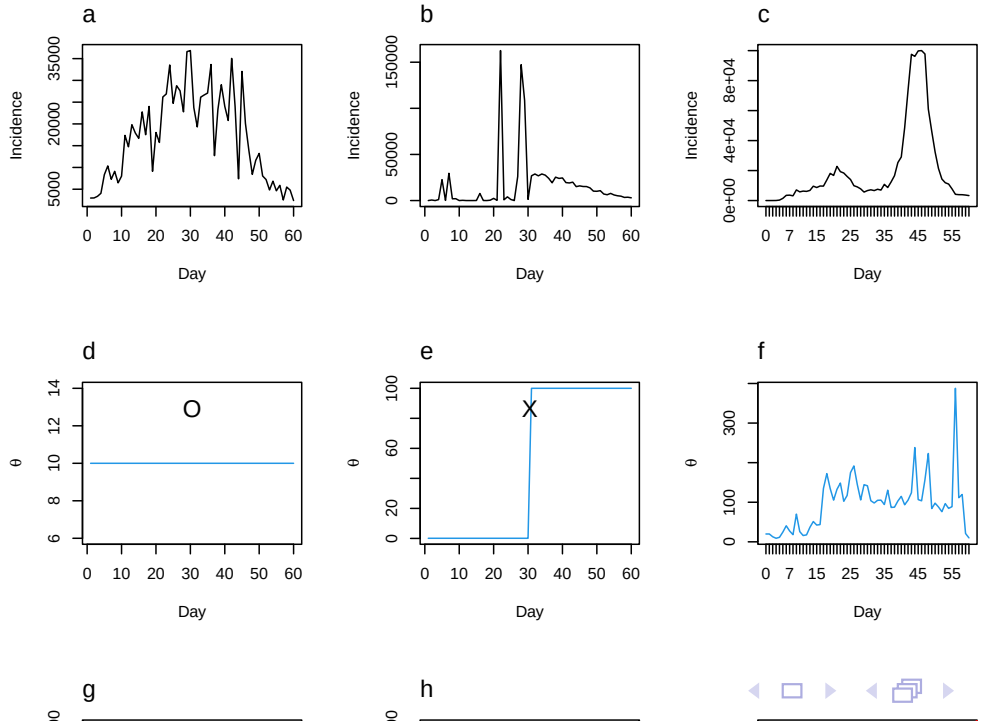
Correspondence simulated and empirical

- We found that the negative binomial/LRT method is robust to differences in population size (for population sizes examined) (the criteria)

Oregon State University

Correspondence simulated and empirical

- We found that the negative binomial/LRT method is robust to differences in population size (for population sizes examined) (the criteria)
- In row one and two of Fig. 1, we illustrated that an increase in θ is associated with decreased variability in simulated incidence time series (same relationship is observable in the empirical time series), with an increase in θ corresponding to a decrease in variability around the trend in incidence



Oregon State University

Applied to counties

- Highly overdispersed incidence patterns were observed more frequently later in time series (consistent)

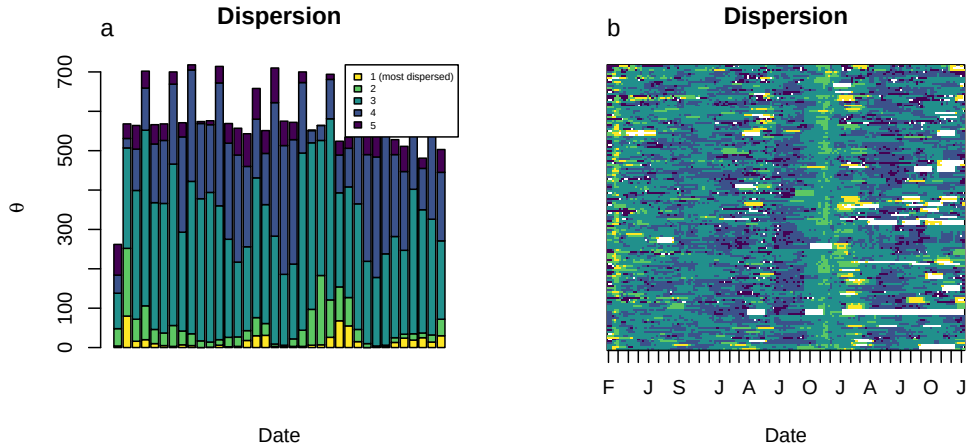
Applied to counties

- Highly overdispersed incidence patterns were observed more frequently later in time series (consistent)
- The most dispersed category in Fig.3 (a) reaches its highest proportion near the end of the timeframe examined, and there are increases in dispersion around the peaks in incidence (Fig. 3.(b,c))

Oregon State University

Applied to counties

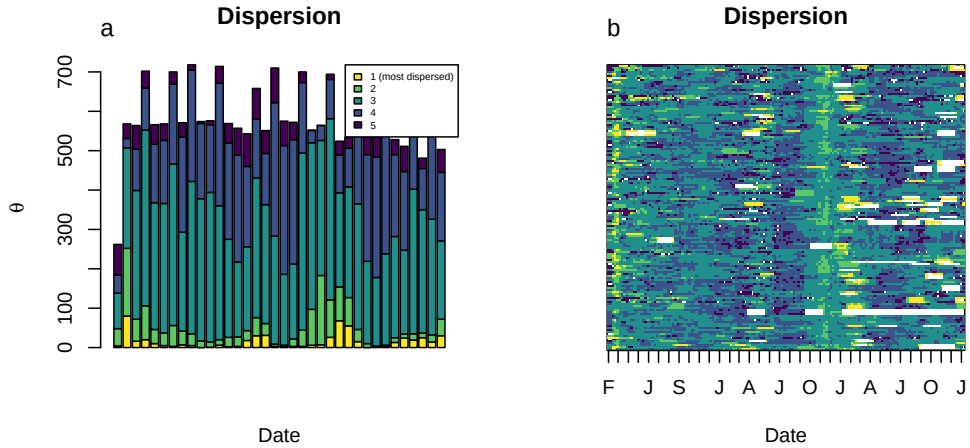
- Highly overdispersed incidence patterns were observed more frequently later in time series (consistent)
- The most dispersed category in Fig.3 (a) reaches its highest proportion near the end of the timeframe examined, and there are increases in dispersion around the peaks in incidence (Fig. 3.(b,c))
- The evidence for a change in θ was observed across many counties (evidenced by concentration of low p-values around peak incidence) (Fig. 3 (d))



Oregon State University

Applied to counties

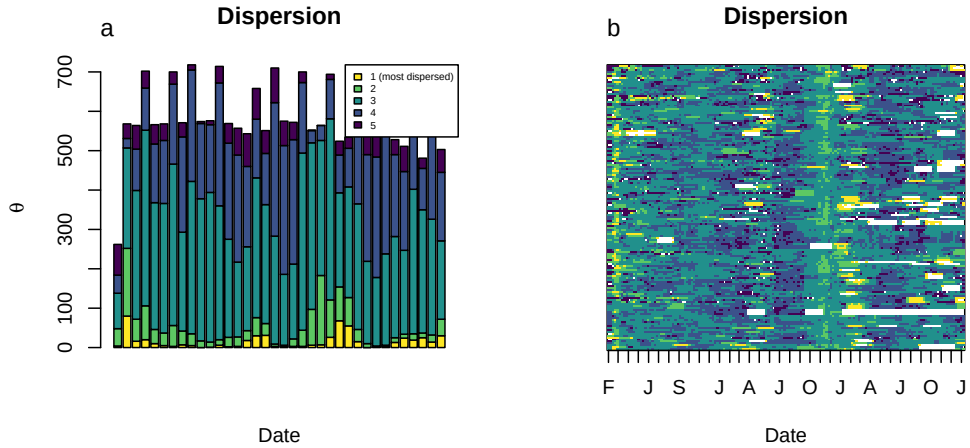
- Highly overdispersed incidence patterns were observed more frequently later in time series (consistent)
- The most dispersed category in Fig.3 (a) reaches its highest proportion near the end of the timeframe examined, and there are increases in dispersion around the peaks in incidence (Fig. 3.(b,c))
- The evidence for a change in θ was observed across many counties (evidenced by concentration of low p-values around peak incidence) (Fig. 3 (d))



Oregon State University

Applied to counties

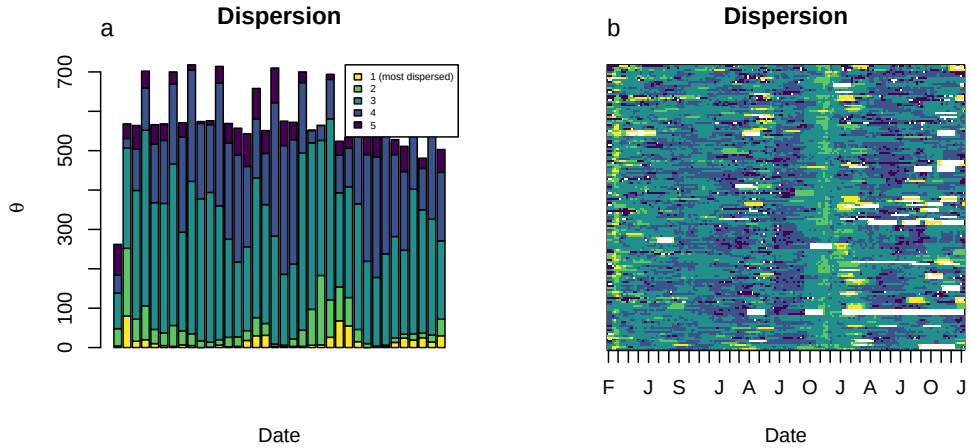
- Highly overdispersed incidence patterns were observed more frequently later in time series (consistent)
- The most dispersed category in Fig.3 (a) reaches its highest proportion near the end of the timeframe examined, and there are increases in dispersion around the peaks in incidence (Fig. 3.(b,c))
- The evidence for a change in θ was observed across many counties (evidenced by concentration of low p-values around peak incidence) (Fig. 3 (d))



Oregon State University

Applied to counties

- Highly overdispersed incidence patterns were observed more frequently later in time series (consistent)
- The most dispersed category in Fig.3 (a) reaches its highest proportion near the end of the timeframe examined, and there are increases in dispersion around the peaks in incidence (Fig. 3.(b,c))
- The evidence for a change in θ was observed across many counties (evidenced by concentration of low p-values around peak incidence) (Fig. 3 (d))



Oregon State University

Table of Contents

Oregon State University

Concluding remarks

- approach to quantify variability in epidemic time series that does not detect artifacts based on population size and incidence (forms)

Oregon State University

Concluding remarks

- approach to quantify variability in epidemic time series that does not detect artifacts based on population size and incidence (forms)
- Methods that use incidence time series are a crucial part of this research area due to the ease of obtaining incidence data, so the timing/geographical allocation of public health resources can be achieved with limited resources

Oregon State University

Concluding remarks

- approach to quantify variability in epidemic time series that does not detect artifacts based on population size and incidence (forms)
- Methods that use incidence time series are a crucial part of this research area due to the ease of obtaining incidence data, so the timing/geographical allocation of public health resources can be achieved with limited resources
- Population-wide disease control approaches are often less effective than those which are targeted to individuals in high-transmission contexts [**lloyd-smith˙superspreading˙2005**] (catalyze the development of more efficient control strategies)

Oregon State University

Concluding remarks

- approach to quantify variability in epidemic time series that does not detect artifacts based on population size and incidence (forms)
- Methods that use incidence time series are a crucial part of this research area due to the ease of obtaining incidence data, so the timing/geographical allocation of public health resources can be achieved with limited resources
- Population-wide disease control approaches are often less effective than those which are targeted to individuals in high-transmission contexts [**lloyd-smith˙superspreading˙2005**] (catalyze the development of more efficient control strategies)
- Our results imply that we can revise our understanding of case count dispersion: dispersion is high at unexpected times (near peak incidence)

Oregon State University

Table of Contents

Oregon State University

Thanks!