



Title	Studies on quantitative assessment of countermeasures during a hepatitis A virus outbreak using renewal equation model
Author(s)	齋藤, 涼平
Description	配架番号 : 2692
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第14950号
Issue Date	2022-03-24
DOI	https://doi.org/10.14943/doctoral.k14950
Doc URL	https://hdl.handle.net/2115/85891
Type	doctoral thesis
File Information	SAITO_Ryohei.pdf



学 位 論 文

Studies on quantitative assessment of countermeasures during a hepatitis A virus outbreak using renewal equation model

(再生方程式モデルによる A 型肝炎のアウトブレイク
対応策の定量的評価に関する研究)

2 0 2 2 年 3 月

北 海 道 大 学

齋 藤 涼 平

学 位 論 文

Studies on quantitative assessment of countermeasures during a hepatitis A virus outbreak using renewal equation model

(再生方程式モデルによる A 型肝炎のアウトブレイク
対応策の定量的評価に関する研究)

2 0 2 2 年 3 月

北 海 道 大 学

齋 藤 涼 平

Index

Journal Publication Index and Conference Presentation

Index: p. 1

Abstract: p. 2

List of Abbreviations and Symbols: p. 5

Background: p. 6

Methods: p. 12

Results: p. 18

Discussion: p. 28

Conclusions: p. 31

Acknowledgement: p. 33

Competing Interests: p. 34

References: p. 35

Appendix: p. 40

Journal Publication Index and Conference Presentation Index

A part of this study was published on following paper:

Ryohei Saito, Akifumi Imamura, Hiroshi Nishiura, Assessing countermeasures during a hepatitis A virus outbreak among men who have sex with men, Theoretical Biology and Medical Modelling, 18:19 (2021)

Abstract

Background

Hepatitis A is an acute infectious disease that frequently causes long-lasting fever, nausea, vomiting, diarrhea, abdominal pain, and strong fatigue, with or without jaundice. Hepatitis A is caused by hepatitis A virus (HAV) and infection of HAV is known to transmit via the fecal-oral route and the ingestion of contaminated food. Since transmission of HAV can occur by direct mucous-to-mucous contact, there is a transition among men who have sex with men (MSM) as a route of transmission of HAV. It has been recognized that MSM sex tourism across multiple countries is a pathway to new outbreaks, and there have been many outbreaks of hepatitis A among MSM. In 2018, Japan experienced a major outbreak, mainly in young men. Recent sporadic cases have occurred among MSM in previous decades, and we can assume that another outbreak occurred in 2018. In this outbreak a particularly substantial increase in disease incidence was observed in Tokyo and Osaka and it was brought under control in about 6 months. Based on the data of confirmed infections in Japan from 2006 to 2019, it can be inferred that the infected individuals in the outbreaks were mainly males, especially in the MSM population, while females did not play the role of infection host and were mainly infected by environmental factors. A campaign was conducted among MSM to reduce the transmission of hepatitis A and our purpose in this study was to evaluate the effectiveness of the intervention by quantitatively assessing the efficacy of this campaign.

Methods

In the Japanese law of Infection Control, hepatitis A is a class 4 notifiable disease and reported within 24 hours of confirmatory diagnosis. The countermeasures against HAV outbreak from 2017 to 2018 were conducted among MSM especially from February 13 to May 5 in 2018: for instance, 1,000 pamphlets were widely distributed on February 29 and an online article were uploaded on March 30, which recorded over 120,000 accesses. Environmental transmission with seasonal variation was described as the linear combination of some trigonometric function with different period. The number of cases newly infected via human-to-human transmission is calculated as convolution of the number of the people infected in the past and the generation time multiplied average secondary cases per one case. Moreover, we assumed that when the first countermeasure was implemented the reproduction number changed. Similarly, when the second countermeasure was conducted at time, the reproduction number changed. This modeling is under assumption that the reproduction number is partly constant and changes at the timing of countermeasures. By taking the sum of environmentally and human-to-human transmitted cases and convoluting it with the incubation period, we obtained the expected number of cases with illness onset. These models are like integral equations called renewal equations, which describe reproduction process of some sort of population. In actual calculations, the model is discretized into a unit time of one week. We assumed that the observed incidence data followed a Poisson

distribution with the expected value of the illness onset, and we obtained the likelihood function. Because parameters can vary with the serial interval, which is assumed to be known in the present study, it is vital to examine how these parameters changes corresponding with the serial parameters. Then we also estimated these parameters for each situation in which the ratio of the standard deviation to the mean (i.e., the coefficient of variation, CV) of the serial interval takes different value as a sensitivity analysis. This study analyzed publicly available data. The datasets used in our study were de-identified and fully anonymized in advance. The analysis of publicly available data without identity information does not require ethical approval.

Results

At first, we specified hyperparameter and the amount of data used in parameter estimation. We can compare goodness of different model by calculation of Akaike Information Criterion (AIC). We tested 9 patterns of model. As a result, the minimum AIC was observed when we adopted simple trigonometric function to describe environmental transmission. Moreover, the human-to-human transmission epidemic was estimated to begin in the 35th week in 2017. Estimated number decomposed into environmental transmission and human-to-human transmission is compared with the observed data, which shows that the human-to-human transmission peaked out when the second countermeasure was implemented. The reproduction number before the interventions, was estimated to range from 2.6 to 3.1, and constant environmental transmission always yielded a greater reproduction number estimate compared with the other cases. The relative risk of transmission at the time of the first countermeasure, which involved the dissemination of pamphlets by non-governmental organizations associated with MSM, ranged from 0.59 to 0.87. Using the 2017 data with trigonometric-function-like environmental transmission, the first decline rate of reproduction number was estimated at 0.72 (95% CI: 0.39-1.04). At the time of the second countermeasure, which used online web articles, the additional relative impact was estimated to range from 0.36 to 0.43. Using the 2017 data with trigonometric function, the second decline rate of reproduction number was estimated at 0.39 (95% CI: 0.27-0.52). Thus, the reproduction number was 0.28 times the baseline; accordingly, the reproduction number fell below one, and the incidence started to decline. It should also be noted that the upper bound of the CI for the reproduction number fell below one following the second countermeasure. We also conducted the same analyses about the alternative model, in which we assumed that the reproduction number declines exponentially. Applying the alternative model with time-dependent effect of interventions to 2015-2017 data, trigonometric function was best-fit, and baseline reproduction number was estimated at 2.00. The two parameters which corresponds with decline rate of reproduction number were estimated to be 0.96 and 1.03, respectively. The AIC value of this model was 1183.6 which was greater than the original model with 1040.4 AIC. A sensitivity analysis was carried out. We examined how sensitive the parameters were to variation in the CV of the serial interval. Except for very small CV values, which would perhaps be unrealistic, the relative risk of transmission remained stable as CV

varied. Thus, the estimated impact of the examined campaigns in terms of preventing hepatitis A transmission was shown to be robust to variation in the generation time, which remained uncertain in our model.

Discussion

Before the first intervention the reproduction number was estimated about 3.0 and then we can guess that without intervention HAV infection spread exponentially. After two interventions the reproduction number decreased below 1.0 which shows that infection changed to decreasing trend after interventions. Reduction of the reproduction number by the second intervention is bigger than one by the first intervention. Therefore, we infer that the second intervention is more effective. On the model we do not consider intrinsic effects: for instance, saturation of infections because of clustering. However, it is likely that extent of transmission did not show a clear clustering pattern. During model analysis we assumed that the empirical data were mixture of renewal process and seasonal transmission. Without knowledge of the timing of infection we cannot know of decline in the reproduction number. Moreover, with respect to seasonal transmission, we assumed that constant hazard was exposed over time. There is room for improvement of model and limitation in modeling: for example, age and sex structure, function setting for the reproduction number, consideration of vaccine effect.

Conclusions

During this epidemic, preventive campaigns were conducted to avoid further transmission in the MSM population. Our study evaluated the effectiveness of these campaigns. Development of a mathematical model and estimation of the parameters using surveillance data captured the epidemic dynamics, quantifying the seasonal variation in environmental transmission. In this study, we objectively shown that the reproduction number abruptly declined following the two examined campaign periods. This finding suggests that publishing online articles may potentially have widespread impact enough to change risky behaviors. Given our study results, it would be valuable to survey the MSM population to investigate which types of input and messages were influential in promoting their behavioral changes.

List of Abbreviations and Symbols

Abbreviation	Explanation
AIC	Akaike Information Criterion
CI	Confidence Interval
COVID-19	Coronavirus Disease of 2019
CV	Coefficient of Variation
DFE	Disease-Free Equilibrium
HAV	Hepatitis A Virus
HIV	Human Immunodeficiency Virus
HPV	Human Papillomavirus
LGBTQ	Lesbian, Gay, Bisexual, Transgender and Queer or Questioning
LSE	Least Square Estimation
MLE	Maximum Likelihood Estimation
MSM	Men who have Sex with Men
NIID	National Institute of Infectious Diseases
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
WHO	World Health Organization

Background

Hepatitis A is known as an acute infectious disease that causes long-lasting fever, nausea, vomiting, diarrhea, abdominal pain, and strong fatigue, with or without jaundice (Franco et al., 2012). Infection of hepatitis A is possibly asymptomatic and most children under the age of six are asymptomatic (Matheny and Kingery, 2012). The disease is likely to cause more severe symptoms in older people than in younger people. An infection of a virus is the path to Hepatitis A: Hepatitis A is caused by hepatitis A virus (HAV) and infection of HAV is known to transmit through the fecal-oral route or the ingestion of contaminated food, especially insufficiently cooked shellfish (Matheny and Kingery, 2012; Bellou et al., 2013). HAV is widespread in the environment and can survive in several food production processes (WHO, 2019). Transmission fluctuates with the seasons, with the main infections occurring in spring and summer (Fares, 2015). While industrialized countries have relative uncommon hepatitis A, HAV is still present globally, spreads sporadically, and causes periodic outbreaks. We still do not know the ecological dynamics of HAV clearly, but the concentration and habitat of shellfish is thought to be dominant (Costantino et al., 2017). An example of a modern outbreak of HAV is the one that occurred in Shanghai in 1988 with 300,000 infections (Cooksley, 2000). After recovering from a natural HAV infection, the infected person achieves life-long immunity (Bennett et al., 2015) and a safe and effective vaccine, inactivated or live-attenuated, is available (WHO, 2012). Efforts to control this infection in public health include providing safe water and food, improving hygiene, promoting hand washing, and immunizing susceptible populations (Fares, 2015). The incubation period of HAV ranges from 14 to 28 days (Bonačić et al., 2015). Elevated serum immunoglobulin M antibodies is used in serological diagnosis (Matheny and Kingery, 2012). No specific treatment for hepatitis A once affected except for routinely supported care is known (Matheny and Kingery, 2012; WHO, 2019).

Since transmission of HAV can occur by direct mucous-to-mucous contact, there is a transition among men who have sex with men (MSM) as a route of transmission of HAV (Matheny and Kingery, 2012). It has been recognized that MSM sex tourism across multiple countries is a pathway to new outbreaks (Watanabe et al., 2019), and there have been many outbreaks of hepatitis A among MSM. Outbreaks have occurred with a time lag, by creating large numbers of infected people in each country as MSM infected people move across countries: from 2015 to 2017 in Taiwan (Chen et al., 2017; Chen et al., 2019; Cheng et al., 2018), from 2016 to 2017 in Europe (Mauro et al., 2019; Ruscher et al., 2019; Alberts et al., 2019; Bradley-Stewart et al., 2019; Marosevic et al., 2019; Sabrià et al., 2019; Petersen et al., 2019; Ndumbi et al., 2018; Boucher et al., 2019; Comelli et al., 2018; Beebeejaun et al., 2017; Freidl et al., 2017; Gozlan et al., 2017; Rodríguez-Tajes et al., 2018; Werber et al., 2017), from 2016-2018 in the United States (Foster et al., 2019; Latash et al., 2017), and in 2018 in Japan (Watanabe et al., 2019; Tanaka et al., 2019). Genome sequencing has clarified Spatial spread, for example from Taiwan and Europe to Japan (Watanabe et al., 2019; Tanaka et al., 2019). A major

outbreak, mainly in young men, occurred in Japan in 2018 (Fig. 1A and B). Recent sporadic cases have occurred among MSM in previous decades (Takechi et al., 2000), and we can assume that another outbreak occurred in 2018. In this outbreak a particularly substantial increase in disease incidence was observed in Tokyo (Fig. 1C) and Osaka (Tanaka et al., 2019) and it was changed under control in about 6 months (Fig. 1A). Based on the data of confirmed infections in Japan from 2006 to 2019, it can be inferred that the infected individuals in the outbreaks were mainly males, especially in the MSM population, while females did not play the role of infection host and were mainly infected by environmental factors (Fig. 1D). A campaign was conducted among MSM to reduce the transmission of hepatitis A. After several campaigns notable decline of infection cases was observed, but relationship between epidemiological quantities like reproduction number and these practical countermeasures was not studied. Our purpose in this study was to evaluate the effectiveness of the intervention by quantitatively assessing the efficacy of this campaign.

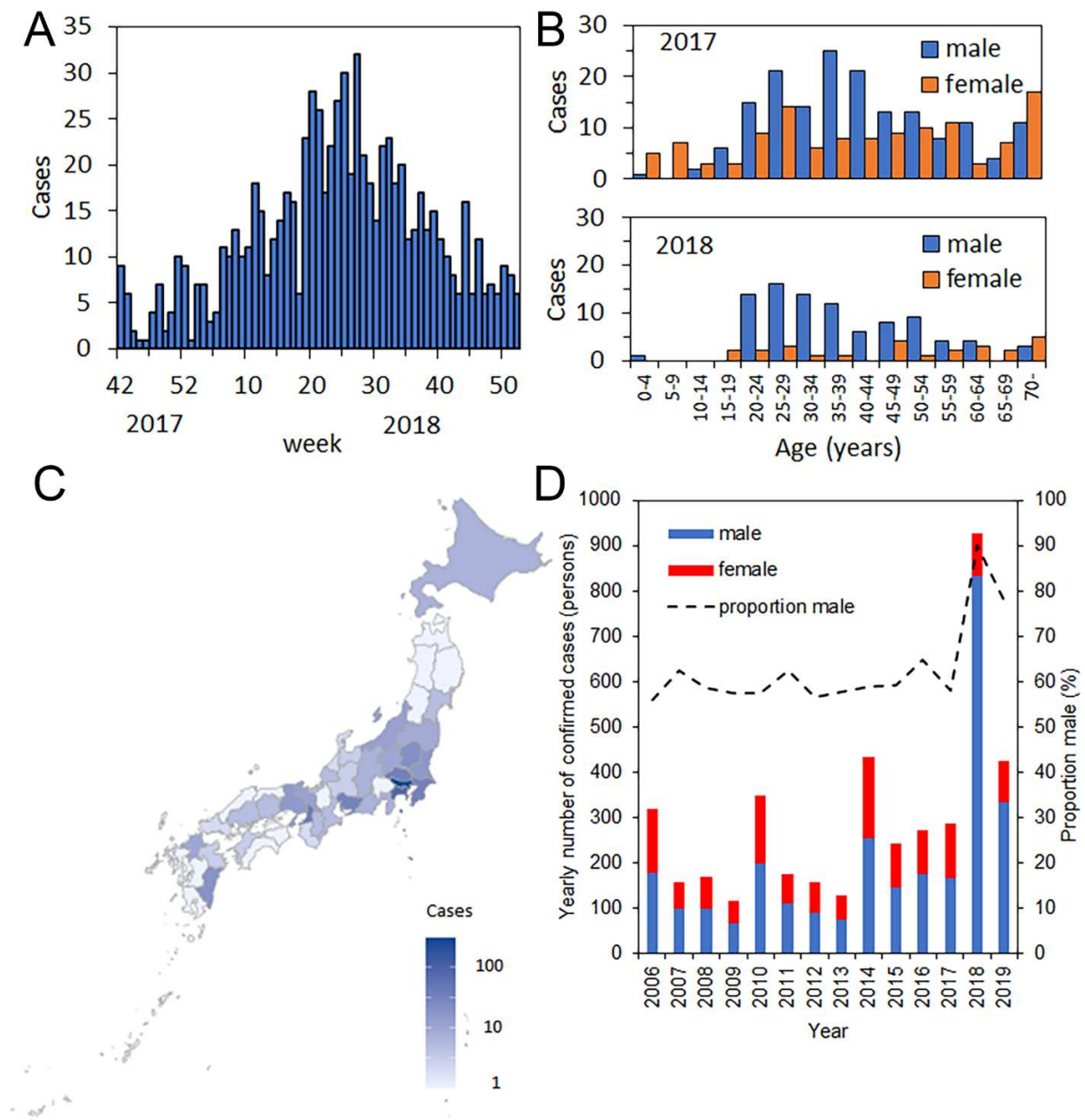


Fig. 1 Epidemic curve of hepatitis A in Japan in 2018 by week, age, and sex distribution. **A** Temporal distribution of the incidence (serologically diagnosed cases) from 2017 to 2018. **B** Age distribution of serologically diagnosed cases. Note that the data for 2017 is the number of cases for the entire year while the data for 2018 is the total through June 26. From the 2017 data, we can read that the age distribution of females is flat, and the age distribution of males is peaking at 35-39 years, and the gender ratio is not so extreme. On the other hand, the 2018 data shows that most cases occur among men aged 20-55. **C** Geographic distribution of serologically diagnosed cases from 2017 to 2018. Most of the cases were in the Kanto region, including central Tokyo, followed by Osaka Prefecture. **D** Sex distribution from 2006 to 2019.

Mathematical model is often used in analyses of dynamics of infectious diseases. The number of empirically observable infected patients can be calculated as a function of time by setting up a mechanism for the transmission of pathogens from humans or other infectious sources to patients. It is generally difficult to express the number of infected people explicitly as a function of time. Therefore, the most common approach is to describe the dynamics of infection as a difference, differential, or integral equation, and to use the solution of the equation as the number of infected people. In practice, infection is considered a transition from an uninfected population to an infected population, and then to a different state of recovery, immunization, or death, thus creating the need to consider multiple populations simultaneously, including susceptible, infected, and recovered populations. The mathematical model constructed in this way is generally called a population dynamics model.

One of the oldest population dynamics models is so-called Fibonacci sequence. If rabbits reproduce two months after birth, and one pair of rabbits increases to F_n pairs after n months, F_n satisfies the difference equation

$$F_{n+2} = F_{n+1} + F_n, F_0 = 0, F_1 = 1. \quad (1)$$

This model describes increase of population as summation of number of adult rabbits multiplied by the constant number 1. This equation has a solution

$$F_n = \frac{\phi^n - (-\phi)^{-n}}{\sqrt{5}}, \quad (2)$$

where $\phi = (1 + \sqrt{5})/2$. Then we can evaluate that the number of rabbits shows asymptotically exponential growth with rate $\log \phi = 0.4812 \dots > 0$. One measure that determines the nature of the solution to a population dynamics model is the growth rate, assuming exponential growth. Depending on whether this is positive or negative, the case of population growth or decline arises.

Similar approach is adopted on the modeling of infectious diseases. Let us write the number of susceptible, infectious, and recovered population at time t as $S(t)$, $I(t)$, $R(t)$, respectively. If we put force of infection β and recovery rate γ , we can construct SIR model (Kermack and McKendrick, 1927):

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases} \quad (3)$$

This is an example of population dynamics model written as nonlinear differential equation. One of our interests is the asymptotic behavior of solution. The number $R_0 = \beta S(0)/\gamma$ is called basic reproduction number, which corresponds to the expected number of secondary infection cases by one infectious individual. If basic reproduction number exceeds over 1 infectious people increase exponentially and vice versa, that is, basic reproduction number determines growth rate of population as asymptotically exponential function of time. The assumption of exponential function is equivalent to the assumption of linear population dynamics model: indeed, we can evaluate basic reproduction number by linearizing the differential equation at the disease-free equilibrium (DFE), in which every individual is assumed to be susceptible (Ma, 2020).

While differential and difference equations are constructed to describe changes in population over short time differences, integral equation models are used to describe the current state through the accumulation of influences from the past. Specifically, let $I(t)$ denote the number of newly infected individuals at time t . If a population of $I(t-s)$ infected in the past by time s results in an average of $\beta(s)$ infections, $I(t)$ satisfies the following equation:

$$I(t) = \int_0^\infty \beta(s)I(t-s)ds \quad (4)$$

This equation is named as renewal equation. If we put $I(t) = 1$ then we obtain newly infected population by one infectious individual, which is equal to basic reproduction number

$$R_0 = \int_0^\infty \beta(s)ds. \quad (5)$$

In this case basic reproduction number is growth rate of infectious population under the assumption of linear infection dynamics, which causes exponential growth.

The second of our interests, which is main subject on our study, is to determine the equation so that the average number of infected people calculated as the solution to the equation best fits the data we have, and to evaluate epidemiological indicators such as the basic reproduction number from that equation. What we are interested in is the momentum of the infection at any given time, and a different approach is adopted from the asymptotic discussion of $t \rightarrow \infty$. Moreover, because using the basic reproduction number requires the assumption that all individuals are susceptible (Heesterbeek and Dietz, 1996), we alternatively use the effective reproduction number R_t , or $R(t)$, the actual number of secondary cases per primary case at time t . The condition $R_t < 1$ suggests that the epidemic is in decline and then under control (Nishiura and Chowell, 2009). We can use the least square estimation (LSE) or the maximum likelihood estimation (MLE) in order to fit the data in possession to the built mathematical model (Ma, 2020).

In the study, we have developed a parsimonious mathematical model for epidemiological investigation of infectious disease dynamics, and since Japan has collected data on all hepatitis A diagnosed cases (NIID, 2019), we fit the model to this data. We also refer to the past HAV transition dynamics (Lin et al., 2019; Van Effelterre et al., 2006; Van Effelterre et al., 2016; Ajelli et al., 2008; de Guimaraens and Codeço, 2005; Jacobse, and Koopman, 2005; Gay, 1996; Hu et al., 1984; Regan et al., 2016) to examine the transition dynamics in Japan.

Methods

Epidemiological data

In the Japanese law of Infection Control, hepatitis A is a class 4 notifiable disease and reported within 24 hours of confirmatory diagnosis. The diagnosis is made by polymerase chain reaction or serologically by elevated immunoglobulin M titer. The data are reported as weekly records of the frequency of cases by the National Institute of Infectious Diseases after anonymizing (NIID, 2019). The countermeasures against HAV outbreak from 2017 to 2018 were conducted among MSM especially from February 13 to May 5 in 2018 (see Table 1 for a detailed chronology): for instance, 1,000 pamphlets were widely distributed on February 29 and an online article were uploaded on March 30, which recorded over 120,000 accesses. Since these two interventions are notably on a large scale, we specially put February 29 as t_1 and March 30 as t_2 , which are to use in construction of mathematical model as timings of change of reproduction number.

Table 1. Chronology of the measure against the outbreak of Hepatitis A among MSM^a in Japan, 2018.

^aMen who have Sex with Men ^bThe time we assumed when the first measure was done. ^cThe time we assumed when the second measure was done.

Calender Time	Time Parameter	Events
13-Feb-18	t_1^b	Submission of data of pamphlet
		Published a web article
		Information via twitter
26-Feb-18		Distributed 305 posters
29-Feb-18		Printed and distributed 1000 pamphlets
9-Mar-18	t_2^c	Distributed 165 posters
9-Mar-18		Published a web article in HIV map (which is accessed by substantial number of men who have sex with men)
10-Mar-18		Distributed 20 posters
12-Mar-18		Distributed 10 posters
17-Mar-18		Put an article on a magazine
30-Mar-18		Published a web article
3-May-18		Distributed 20 pamphlets
5-May-18		Distributed (i) 165 and (ii) 1815 pamphlets for (i) delivery health project, Shinjuku 2-chome and bar and (ii) Tokyo Rainbow Pride Parade (for LGBTQ) in Yoyogi Park, respectively

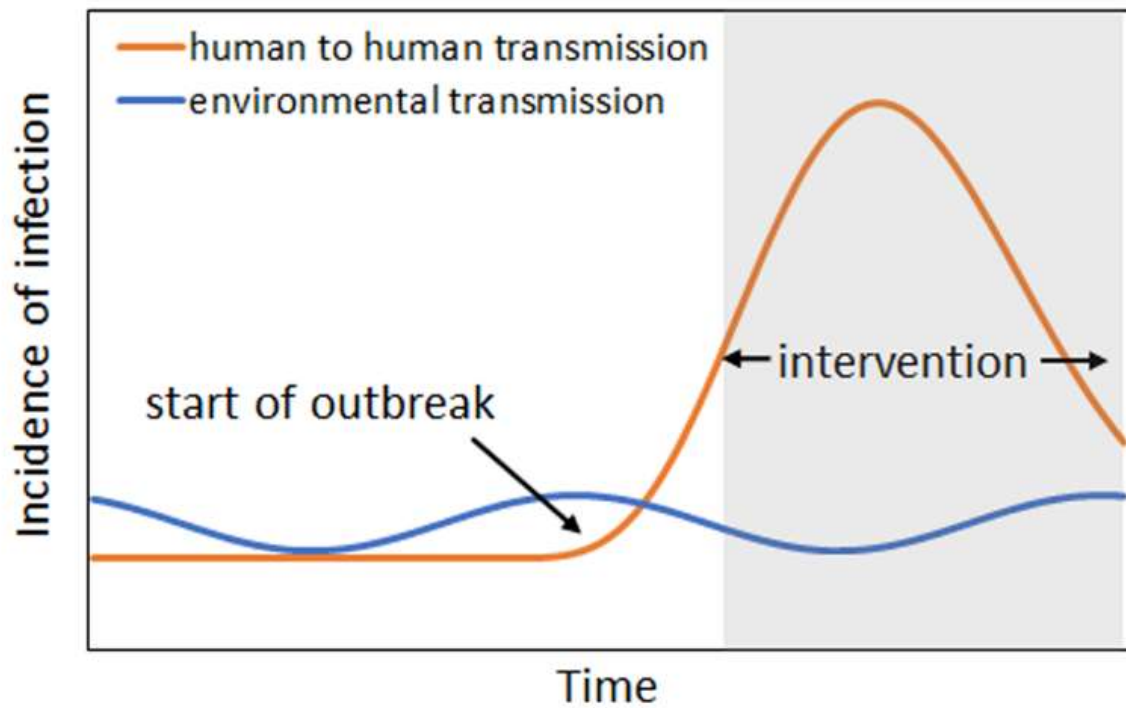


Fig. 2 The patterns of the epidemic curve in which the observed epidemic curve was assumed to be decomposed into environmental transmission and human-to-human transmission. Environmental transmission, which varies seasonally, corresponds to a periodic curve. The other curve represents human-to-human transmissions among men who have sex with men.

Mathematical model

Environmental transmission with seasonal variation was described as the linear combination of some trigonometric function with different period. The model is shown below:

$$a(t) = a_0 + \sum_{k=1}^n \sin\left(\frac{2\pi kt}{52} + b_k\right), \quad (6)$$

where $a(t)$ is the number of newly infected cases caused by environmental transmission at time t (weeks), n is the number of terms in the series of trigonometric function, a_0 and $a_k, b_k (k = 1, \dots, n)$ are the parameters estimated by data-fitting.

Let $j(t)$ be the number of cases newly infected via human-to-human transmission at time t . Each $j(t)$ is calculated as convolution of the number of the people infected in the past and the generation time multiplied average secondary cases per one case R_1 :

$$j(t) = R_1 \int_0^t j(t-s)g(s)ds. \quad (7)$$

Here $g(s)$ is the discretized probability density function of the generation time at time $G(s) - G(s-1)$, where $G(s)$ is the cumulative distribution function of the generation time. The probability distribution is assumed to be a lognormal distribution with a median of 27.5 days and a standard deviation of 4 days (Bonačić et al., 2015). Moreover, we assumed that when the first countermeasure was implemented at time t_1 the reproduction number changed from R_1 to $R_1\varepsilon_1 (0 < \varepsilon_1 < 1)$. Therefore, the renewal equation after time t_1 changed to

$$j(t) = R_1\varepsilon_1 \int_0^t j(t-s)g(s)ds. \quad (8)$$

Similarly, when the second countermeasure was conducted at time t_2 , the reproduction number changed from $R_1\varepsilon_1$ to $R_1\varepsilon_1\varepsilon_2 (0 < \varepsilon_2 < 1)$. Therefore, the renewal equation changed to

$$j(t) = R_1\varepsilon_1\varepsilon_2 \int_0^t j(t-s)g(s)ds. \quad (9)$$

This modeling is under assumption that the reproduction number is partly constant and changes at the timing of countermeasures t_1, t_2 . We also examined the alternative assumption in which the intervention effects were time-dependent: considering $R_1\varepsilon_1^{t-t_1}$ rather than $R_1\varepsilon_1$ and $R_1\varepsilon_1^{t_2-t_1}\varepsilon_2^{t-t_2}$ rather than $R_1\varepsilon_1\varepsilon_2$.

By taking the sum of environmentally and human-to-human transmitted cases $a(t), j(t)$ and convoluting it with the incubation period $f(t)$, we obtained the expected number of cases with illness onset $\mathbb{E}[c_t]$ at time t :

$$\mathbb{E}[c_t] = \int_0^t (j(t-s) + a(t-s))f(s)ds. \quad (10)$$

Here, $f(s)$ is the probability density function of the incubation period with a median of 28 days and a standard deviation of 9 days (Bonačić et al., 2015). In actual calculations, the above model is discretized into a unit time of one week. The integral $\int$, which is a continuous sum, is treated as a discrete sum $\sum$. We assumed that the observed incidence data followed a Poisson distribution with the expected value $\mathbb{E}[c_t]$ and we obtained the likelihood function as

$$L(t_\theta, R_1, \varepsilon_1, \varepsilon_2 | \mathbf{x}_t) = \prod_t \frac{\mathbb{E}[c_t]^{x_t} \exp(-\mathbb{E}[c_t])}{x_t!}, \quad (11)$$

given the observed number of cases x_t in week t , where t_θ is the week in which the first human-to-human transmission took place, that is, we assumed that there were only environmental transmissions prior to t_θ .

Because parameters $\varepsilon_1, \varepsilon_2$ can vary with the serial interval, which is assumed to be known in the present study, it is vital to examine how these parameters changes corresponding with the serial parameters. Then we also estimated these parameters for each situation in which the ratio of the standard deviation to the mean (i.e., the coefficient of variation, CV) of the serial interval takes different value as a sensitivity analysis. We considered the following alternative model as part of sensitivity analysis, accounting for potential dependence between environmental and human-to-human transmissions, i.e.,

$$j(t) = R(t) \int_0^t (j(t-s) + a(t-s))g(s)ds, \quad (12)$$

where $R(t)$ stands for the effective reproduction number at time t and then,

$$\mathbb{E}[c_t] = \int_0^t j(t-s)f(s)ds. \quad (13)$$

In this modeling the following assumptions were made. First, we assumed that the generation time followed a lognormal distribution. Second, each countermeasure was assumed to abruptly influence the infection dynamics. Third, we assumed that the two interventions independently influenced the epidemic dynamics. This assumption means that the combined impact could be modeled as the product of ε_1 and ε_2 . Fourth, the epidemic dynamics were assumed to be a simple

additive function of environmental transmission and human-to-human transmission, and the former was assumed to be approximated by a periodic function. Finally, we assumed that all infectious people were diagnosed and reported.

Ethical considerations

This study analyzed publicly available data. The datasets used in our study were de-identified and fully anonymized in advance. The analysis of publicly available data without identity information does not require ethical approval.

Results

Hyperparameter Selection

At first, we specified hyperparameter n and the amount of data used in parameter estimation. We can more precisely fit the data to the model as n increases. However, on the other hand, too many trigonometric terms invoke overfitting problem. We can compare goodness of different model by calculation of Akaike Information Criterion (AIC). We tested 9 patterns of model: for n varying from 0 to 2 and data in 2015-2017, 2016-2017, and 2017. As a result, the minimum AIC was observed when we adopted $n = 1$. Hence, we decided to put $n = 1$ in the model combining environmental transmission and human-to-human transmission. In Figure 3A the prediction using 2017 data as learning data and the observed data is compared, and this figure shows that most the observed data were contained within the 95% confidence intervals (CI) of the predictions. Moreover, the human-to-human transmission epidemic was estimated to begin in the 35th week in 2017. Estimated number decomposed into environmental transmission and human-to-human transmission is compared with the observed data in Figure 3B, which shows that the human-to-human transmission peaked out when the second countermeasure was implemented.

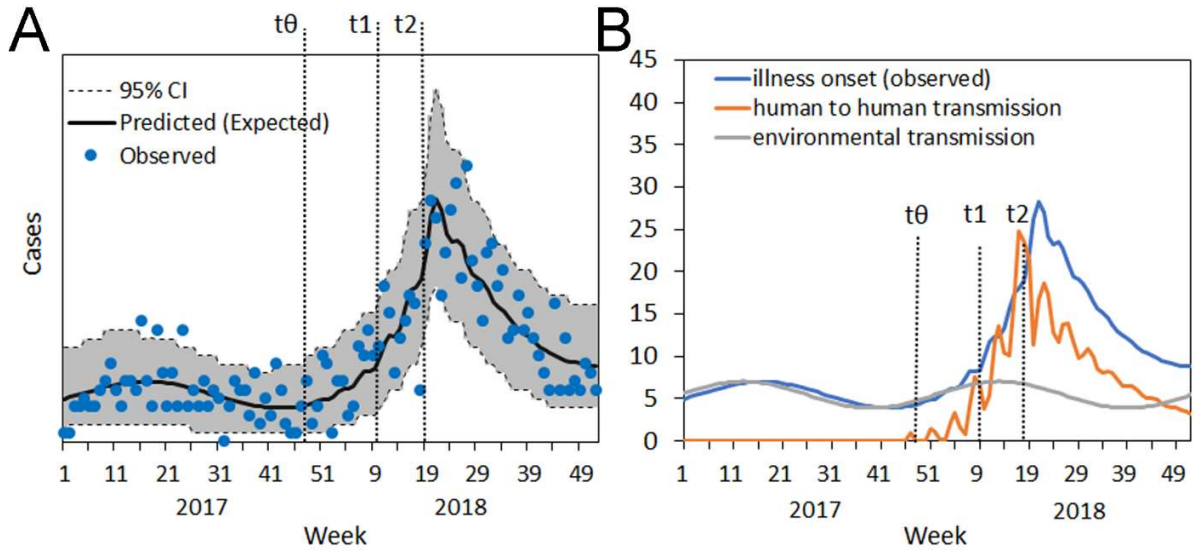


Fig. 3 The epidemic curve by fitted model. **A** Predicted and observed weekly incidence. The solid black line depicts the expected weekly incidence of hepatitis A, and the blue dots represent the empirical data. The gray shaded area is the 95% confidence interval. We derived this interval by the parametric bootstrap method. The times at which (i) human-to-human transmission began to increase, (ii) the first countermeasure was implemented, and (iii) the second countermeasure was implemented, was represented by three vertical dotted lines from left to right. **B** Decomposed temporal distribution. The gray wave shows the predicted incidence of environmental transmission, the orange curve represents the predicted incidence of human-to-human transmission, and the blue line represents the summation of these two mechanisms which was calculated as a convolution of the mixed incidence and the incubation period.

Table 2. Results of parameter estimation and 95% CI, time of outbreak, number of parameters, and AIC^a for each model assumption. It was assumed that the outbreak started in week 35, 2017. ^aAkaike Information Criteria ^breproduction number ^ccoefficient of the first measure ^dcoefficient of the second measure ^eestimated time of invention of the first case of Hepatitis A by MSM.

Model assumption	R_1^b	ε_1^c	ε_2^d	Number of parameters	AIC
Constant, 17	3.05 (2.63,3.47)	0.60 (0.38,0.82)	0.41 (0.30,0.52)	5	566.0
Sine curve, 17	2.80 (2.32,3.29)	0.72 (0.39,1.04)	0.39 (0.27,0.52)	7	554.6
Second-order trigonometric, 17	2.80 (2.32,3.29)	0.72 (0.39,1.04)	0.39 (0.27,0.52)	9	558.6
Constant, 16-17	3.07 (2.66,3.49)	0.59 (0.37,0.80)	0.42 (0.30,0.53)	5	832.8
Sine curve, 16-17	2.74 (2.24,3.24)	0.76 (0.40,1.12)	0.38 (0.26,0.51)	7	797.8
Second-order trigonometric, 16-17	2.60 (2.07,3.13)	0.87 (0.42,1.33)	0.35 (0.23,0.47)	9	798.3
Constant, 15-17	3.13 (2.71,3.53)	0.57 (0.37,0.77)	0.43 (0.32,0.54)	5	1079.1
Sine curve, 15-17	2.82 (2.35,3.30)	0.72 (0.40,1.03)	0.39 (0.27,0.51)	7	1040.4
Second-order trigonometric, 15-17	2.68 (2.18,3.19)	0.82 (0.42,1.27)	0.36 (0.24,0.47)	9	1038.0

Parameter Estimation and Evaluation of Reproduction Number

Table 2 shows the summary of the estimated parameters along with their 95% CIs. Seasonality (i.e., $n = 1, 2$) was detected rather than constant baseline (i.e., $n = 0$). The AICs of the models can be compared when models adopt the same learning data, the same amounts of empirical data across the time periods. We now present the results with the 2017 learning data because we focused human-to-human transmission in the data in the year corresponding to the outbreak. We estimated the reproduction number before the interventions, R_1 , from 2.6 to 3.1, and R_1 was estimated larger in $n = 0$ rather than in $n = 1$ or $n = 2$. The relative risk of transmission by the first countermeasure ranged from 0.59 to 0.87. We estimated ε_1 as 0.72 (95% CI: 0.39-1.04) using the 2017 data with $n = 1$. Additionally, we estimated the relative risk by the second countermeasure and the result was from 0.36 to 0.43. We estimated ε_2 was estimated at 0.39 (95% CI: 0.27-0.52) using the 2017 data with $n = 1$. After two countermeasures the reproduction number was $0.72 \times 0.39 = 0.28$ times the baseline and the reproduction number fell below one. We simultaneously observed the incidence started to decline, and after the second countermeasure even the upper bound of the CI for the reproduction number fell below one (Fig. 4).

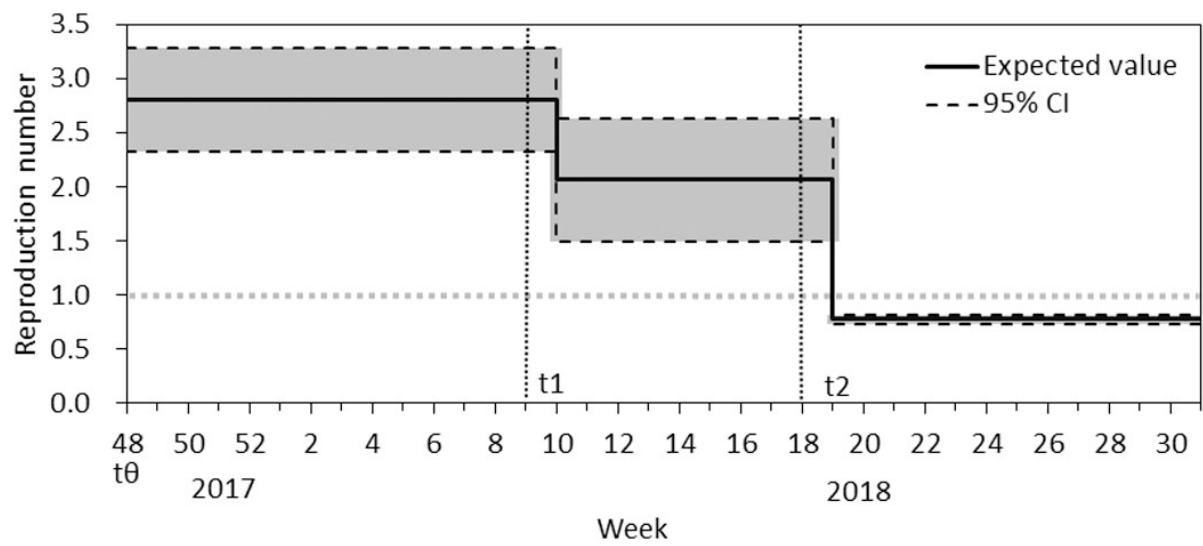


Fig. 4 Results of estimation of effective reproduction numbers. The solid line represents the reproduction number, and the gray shaded area represents the 95% CI. The reproduction number decreased when the first and second countermeasures were implemented, and it fell below one after the second countermeasure.

Sensitivity Analysis

Moreover, we conducted a sensitivity analysis. Our target was the extent of change in ε_1 and ε_2 corresponding to variation in the CV of the serial interval (Fig. 5). The relative risk of transmission was basically stable nevertheless the CV changed. The parameters do indeed show changes when the CV is very small, but the CV in that case is so small that it is safe to decide that it is not realistic. In our model the generation time which we used has uncertainty, but it was shown that the estimation of the estimated impact of the countermeasures against HAV was robust to it.

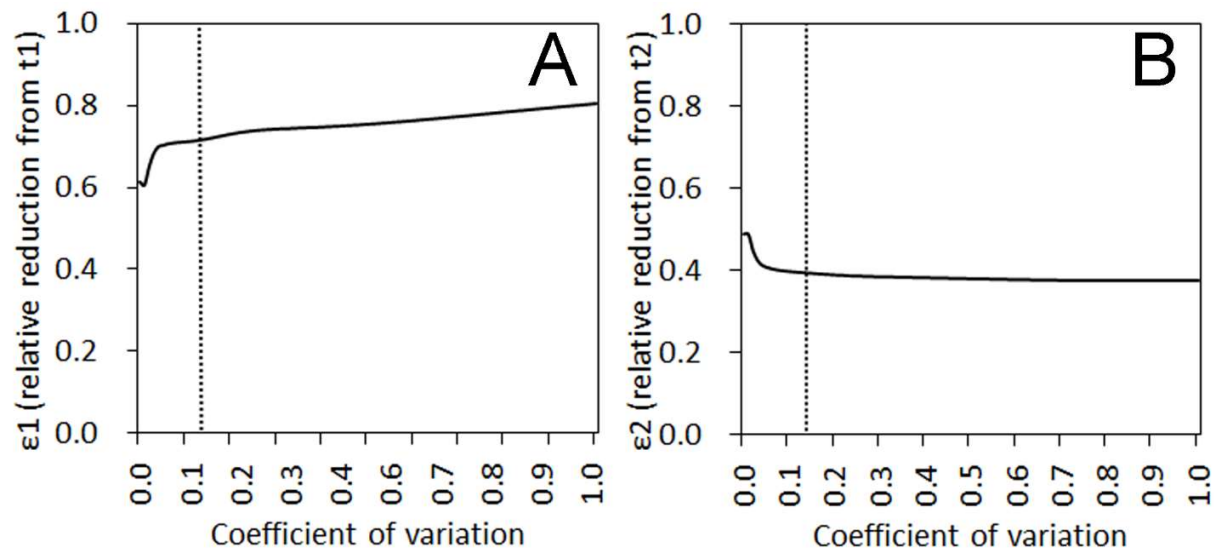


Fig. 5 Sensitivity of the relative risk of transmission to serial interval variations. We fixed the median of the serial interval at 27.5 days (Bonačić et al., 2015). We varied the variance value for variation of the CV from 0 to 1. The value which we used in the actual estimation was shown by the vertical line.

Trial on Another Models and Computations

In addition to the mathematical models that we mainly discussed in this paper we have also proposed models with different structures. However, these did not give good estimation results.

First, we assumed that effect of interventions is dependent on time, especially the reproduction number is exponential function of time whose bases are unknown parameters $\varepsilon_1, \varepsilon_2$. Applying the alternative model with time-dependent effect of interventions to 2015-2017 data, sine function (corresponding to $n = 1$) was best-fit, and baseline R_1 was estimated at 2.00. Parameters $\varepsilon_1, \varepsilon_2$ were estimated to be 0.96 and 1.03, respectively. The AIC value of this model was 1183.6 which was greater than the original model with AIC=1040.4 (Table 2).

Second, we thought human-to-human and environmental transmission simultaneously. We composed the alternative model under following thought: summation of infection via human-to-human transmission and environmental transmission from past of time s yields new illness onset after time x , which was expressed as following integral equation:

$$\mathbb{E}[c_t] = R_t \iint (f(t-s-x) + a(t-s-x))g(s)f(x)dsdx. \quad (14)$$

The alternative model yielded subcritical value of R_1 , which is understandable for continued full involvement of environmentally infected individuals in the chains of transmission (but $R_1 < 1$ was unlikely if the epidemic was induced mainly by contact in the population of MSM). For instance, using the data from 2015-2017 and adopting a constant baseline, R_1 was estimated at 0.83 (95% CI: 0.51-1.14). AIC values of this alternative approach was always greater than those of original model.

Third, we can guess that differential equation model is also effective. On modeling using differential equation we can introduce age or infectious age structure then a partial differential equation (PDE) is generated. If we consider age structure, we obtain a PDE in two variables:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \tau} + \frac{\partial}{\partial a} \right) i(t, \tau, a) = -(\mu(t, a) + f(\tau))i(t, \tau, a), \quad (15)$$

Where a is age, τ is infectious age, μ is force of mortality, and f is force of infection. If we consider just a short time interval in which demographic change can be ignored, we can omit μ from the equation. If we additionally erase age structure, we have an ordinary differential equation (ODE) model. However, this idea was rejected because of insufficient data. On previous study about human papillomavirus (HPV) we put following two PDE models:

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) S(t, a) = -(\mu(t, a) + \lambda(t, a))S(t, a) \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \tau} + \frac{\partial}{\partial a} \right) i(t, \tau, a) = -(\mu(t, a) + f(t, a))i(t, \tau, a) \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) C(t, a) = -(\mu(t, a) + \mu_c(t, a))C(t, a) + \int_0^a f(\tau)i(t, \tau, a)d\tau, \end{array} \right. \quad (16)$$

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) S(t, a) = (\mu(t, a) + \lambda(t, a))S(t, a) \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) I(t, a) = \lambda(t, a)S(t, a) - (\mu(t, a) + \eta(t, a))I(t, a) + \sigma(t, a)N(t, a) \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) N(t, a) = \eta(t, a)I(t, a) - (\mu(t, a) + \rho(t, a) + \sigma(t, a))N(t, a) \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) C(t, a) = \rho(t, a)N(t, a) - (\mu(t, a) + \mu_c(t, a))C(t, a), \end{array} \right. \quad (17)$$

Where S, I, N, C represents susceptible, infectious, cervical intraepithelial neoplasia, and cervical cancer, respectively. Moreover, force of infection, additional force of mortality by cancer, procession to neoplasia, procession to cancer, and recovery from neoplasia are denoted as $\lambda, \mu_c, \eta, \rho, \sigma$ respectively. However, on practical analysis we significantly underestimated the prevalence rate of HPV because of lack of illness onset data. Structured population dynamics model is too complex to yield reliable estimation and available data in HAV is fewer than that in HPV.

Other Trials on Data Analysis Proccession

Estimation and additional analysis were conducted by running R program pasted in Appendix. We also tried various computation methods.

For instance, we changed width of time unit. On the program in Appendix time unit is set as one week because illness onset data are reported as number per week. However, we can set time unit arbitrarily and we tried to estimate parameters by models with short-or-long time unit. In the case of short time unit, we failed to estimate parameters. It is quite natural because only data per week can not specify the dynamics in more detailed time scale than week: there can be many scenarios which concludes the same illness onset value as the data we have. On the other hand, with wider time unit we could achieve parameter estimation itself. However, evaluation of AIC shows that wider time units than one week are worse settings. We can guess that wider time unit requires summing up data roughly, which derives notable decrease of information.

Moreover, variations of numerical procession can be considered. The integral term of renewal equation is named as convolution mathematically, therefore we can solve renewal equation by two methods: matrix multiplication consisted of discretized function values and numerical integration. The former method supplied excel computation and the latter ended in divergence of optimization of likelihood function.

Additionally, hyperparameters t_θ and initial values were changed. Intervals of change covered all realistic value; for instance, t_θ moved in whole week in the year when the outbreak occurred, and initial values moved beyond all confirmed literature value. Whether likelihood maximization converges or not was numerically sensitive and almost all value resulted in divergence, and successful estimation of parameters are achieved in quite limited intervals enough to make the same conclusion despite of variation of hyperparameters. The program in Appendix reports one of the hyperparameters in which optimization process ended up with convergence.

Discussion

Prior studies have examined the rapid increase in HAV transmission among MSM in Japan in 2018 (Tanaka et al., 2019). The policy for dealing with this outbreak was to prevent further spread among the MSM population and the effectiveness of these campaigns was evaluated in this study. We used surveillance data to estimate parameters, which provided a quantification of the seasonal transition due to the structure of the transmission pathway, separating human-human and environmental. The reproduction number ranged from 2.6 to 3.1 before the countermeasures. The interventions by the campaigns caused decrease of this value. However, the reproduction number remained above one after the first countermeasure. The second countermeasure made R_t fell below one. Both intrinsic and extrinsic factors can change effective reproduction number R_t , but the present study suggests the impact of the interventions for situation awareness of HAV outbreak among MSM. Vaccine countermeasures were not considered this time, but vaccines were required to control recent outbreaks of HAV in other countries; for example, the United States (Foster et al., 2019; Friedman et al., 2000) and Taiwan (Lin et al., 2019).

There indeed is the possibly of intrinsic effects, for example, saturation of infections because of clustering, cannot be excluded. Therefore, the decline in R_t based on extrinsic factors is estimated and then our result is conditional. However, it is likely that risky contact was restricted by behavioral changes after the countermeasures because any clear clustering pattern was not found and the size of MSM population is thought to be substantial. We focused on the two campaigns and the latter campaign, publishing an online article, had a greater impact, that is, that reduced the reproduction number more intensively. Our result suggests that the MSM population should be investigated in order to know more effective messages to promote their behavioral changes.

The risk of HAV infection is continuously lies in the MSM population in Japan and travelers visiting Japan bring such risk. Thus, possible preventive vaccination against HAV should be considered even in the situation that the examined campaigns were successful, and this time period did not experience any vaccination. It would be one of future studies to examine acceptance of and demand for this vaccine.

The possibility that the decline in the reproduction number assessed in this study was caused by factors other than large-scale interventions should also be examined. Although there are no previous studies on the factors that led to the decrease in outbreaks in Japan in 2018, the subject of this study, the results of studies on promotional strategies for HIV (LaCroix et al., 2014) and studies assessing the effects of measures on past HAV outbreaks in Berlin (Zimmermann et al., 2021) suggest the effectiveness and necessity of alerting the MSM community. If this countermeasure caused behavioral changes in MSM as in similar sexually transmitted diseases and past outbreak cases, then the calculated reduction in the reproduction number in this study

provides a quantitative assessment of the effectiveness of the countermeasure. Although it is possible that the behavioral changes in MSM were caused by factors other than countermeasure that naturally increased interest, this possibility is unlikely because the acute decline in the reproduction number occurred in harmony with the timing of the large-scale intervention. If interest in nature had increased, behavioral change should have proceeded slowly, and as a result, the decrease in the reproduction number should have continued to decrease continuously with time. However, modeling based on such an assumption, i.e., an exponential decrease in the reproduction number, is numerically divergent. In addition, the fact that the literature values for the incubation period and the generation renewal period are nearly equal and consistent with the gap between the time of intervention and the peak of infection shown in Fig. 3 reinforces the influence of the intervention on behavior change. Web articles and other alerts provide basic information about HAV, prevention methods, and vaccination recommendations. The vaccination policy in Japan at the time of 2018 was three subcutaneous injections: the first, one month, and six months later, and the time lag until immunity was acquired was too long to account for the rapid decline in infection during the outbreak.

We modeled observed data about HAV infection under the assumption that a renewal processes of human-to-human transmission and a series of periodic function of environmental transmission were mixed. We cannot obtain likelihood function in closed form by this approach, but numerical computation enabled us to maximize the likelihood function, modeling the illness onset data as a convolution of incidence of infection and probability density function of incubation period. We can explicitly evaluate the declines of the reproduction number only after knowing the likely timing of infection. Compared with assumption of a constant risk over time, seasonally fluctuating exposures were thought to be plausible with respect to environmental transmissions. This implies that there is seasonal variation of risks. However, the control period was assumed that there was no outbreak among MSM. In this sense, there is room for improvement. As written in the result section, we built an alternative model with consideration dependence between environmental and human-to-human transmissions, but AIC values of such models were greater than original models and the resulting parameters could not fully justified.

We must mention several limitations. First, the model can contain sexual contact or age structure. Since this outbreak was caused by HAV as a sexually transmitted disease, it was necessary to pay special attention to young males, but we had to give up due to insufficient sample size. Second, the effectiveness can be indicated directly by monitoring contact patterns itself. The setting of the reproduction number as a step function potentially reflects contact patterns over time, but we would be able to obtain more clear evidence in the effectiveness of countermeasures by observing diminishing patterns of risky contact. Third, our analysis depends on reported surveillance data. Unreported cases should be considered for more rigorous fitting. However, it should be added that underreporting of HAV infection is unlikely to be very frequent. Fourth, if we trace genome and travel-history data we can directly classify the data. In that sense

our model assumption of a mixture of environmental and human-to-human transmission can be improved. These limitations are mainly due to the absence of data to allow such modeling.

Although our analysis has these limitations, the timing of countermeasures against HAV for MSM population in Japan can be thought to be consistent with abrupt declines in R_t . Our finding suggests that risky behaviors on HAV transmission among MSM population can be sufficiently changed by publishing online articles.

Conclusions

Overview of our study

On the present study the epidemic of HAV among adult men in 2018 in Japan and the preventive campaigns to stop further transmission in the MSM population was examined. Some mathematical models were composed and goodness of fitting between model governed by parameters and available data. We discussed dynamics of infection based on the estimated parameters.

New findings from this study

- The estimate of reproduction number during outbreak

The reproduction number before the countermeasures ranged from 2.6 to 3.1. Without any intervention HAV virus transmitted from one infectious individual to this number of people on average. Epidemic curve regenerated from mathematical model shows that we experienced exponential growth of infection before the interventions. Such a quantitative evaluation of the reproduction number on HAV outbreak in Japan, 2018, by human-to-human transmission separated from environmental factors were revealed from this study.

- The estimate of decline ratio of reproduction number by each outbreak

The reproduction number slightly declined after the first countermeasure, but it remained over one, the threshold of increase or decrease of infection. When the second countermeasure was conducted the reproduction number changed below one. These results of estimation gave the first evaluation of effectiveness of countermeasures against HPV in the sense of the reproduction number.

- Proposition of a new framework for assessing the effectiveness of infection prevention measures

There are many countermeasures against infectious diseases in public health and their effectiveness should be under consideration. We demonstrated a method of evaluation of such effectiveness by considering theoretical dynamics underlying the data that we can directly observe and by estimating epidemiological indicators.

- Usefulness of efforts for reduction of risky behaviors by MSM

Result of parameter estimation suggests that risky behaviors among MSM during HAV outbreak were suppressed by practically implemented countermeasures, especially publication of web articles with frequently access. Our method gives such a suggestion which countermeasures should be selected under limited resources in, for instance, time and budget.

Future studies

We showed possibility of quantitative assessment of countermeasure under context of theoretical epidemiology. However, on this study we considered simple situation without, for example, population structure and vaccine effect. Then we can think extended models whose range goes into various infectious diseases and various situation. For instance, when we consider venereal diseases sex and age structures are essential. Additionally, one of major countermeasures is vaccination. In the outbreak of COVID-19 vaccination plays an important role and as another noteworthy factor spatial or geographical analyses can be under consideration. In order to assess effectiveness of policies to restrict the amount of going out or lockdowns that prohibit travel between municipalities spatially structured models are likely to be useful.

On the other hand, expanded models generally require more data. On this study we had to verify possibility of computation of each model by running the calculations after building the model one by one. The issue we are particularly interested in is the theoretical assessment of the sensitivity and uncertainty associated with mathematical modeling, to assess in advance the complexity of the model to be employed and to ensure the reliability of the data analysis by mathematical modeling in infectious diseases.

Acknowledgement

I would like to thank my supervisor Hiroshi Nishiura for his extraordinary support and patient instruction. Moreover, I am grateful to all laboratory members for their kind help in terms of technique or knowledge.

This study was mainly supported by Health and Labor Sciences Research Grant (H29-AIDS-General-006; PI: Akifumi Imamura). Hiroshi Nishiura received funding from Health and Labor Sciences Research Grants (19HB1001, 19HA1003, 20CA2024, 20HA2007 and 21HB1002), the Japan Agency for Medical Research and Development (AMED; JP19fk0108104, JP20fk0108140 and JP20fk0108535s0101), the Japan Society for the Promotion of Science (JSPS) KAKENHI (17H04701 and 21H03198), the Japan Science and Technology Agency (JST) CREST program (JPMJCR1413), and the SICORP (e-ASIA) program (JPMJSC20U3 and JPMJSC2105). The funders were not involved in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the work for publication.

Competing Interests

The author declares no conflict of interest.

References

- Ajelli M., Iannelli M., Manfredi P., and Ciofi degli Atti ML. (2008). Basic mathematical models for the temporal dynamics of HAV in medium-endemicity Italian areas. *Vaccine*. 26, 1697–707.
- Alberts C.J., Boyd A., Bruisten S.M., Heijman T., Hogewoning A., Rooijen V., Siedenburg E., and Sonder G.J.B. (2019). Hepatitis A incidence, seroprevalence, and vaccination decision among MSM in Amsterdam, the Netherlands. *Vaccine*. 37, 2849–56.
- Beebeejaun K., Degala S., Balogun K., Simms I., Woodhall S.C., Heinsbroek E., Crook P.D., Kar-Purkayastha I., Treacy J., Wedgwood K., et al. (2017). Outbreak of hepatitis A associated with men who have sex with men (MSM), England, July 2016 to January 2017. *Euro Surveill*. 22, 30454.
- Bellou M., Kokkinos P., and Vantarakis A. (2013) Shellfish-borne viral outbreaks: a systematic review. *Food Environ. Virol*. 5, 13–23.
- Bennett J.E., Dolin R., and Blaser M.J., eds. (2015). *Mandell, Douglas, and Bennett's principles and practice of infectious diseases*. 8th ed. New York: Saunders. <https://www.sciencedirect.com/book/9781455748013/mandell-douglas-and-bennetts-principles-and-practice-of-infectious-diseases>. Accessed 26 May 2021.
- Bonačić Marinović A., Swaan C., van Steenbergen J., and Kretzschmar M. (2015). Quantifying reporting timeliness to improve outbreak control. *Emerg. Infect. Dis*. 21, 209–16.
- Boucher A., Meybeck A., Alidjinou K., Huleux T., Viget N., Baclet V., Valette M., Alcaraz I., Sauser E., Bocket., et al. (2019). Clinical and virological features of acute hepatitis A during an ongoing outbreak among men who have sex with men in the North of France. *Sex Transm. Infect*. 95, 75–7.
- Bradley-Stewart A., Smith-Palmer A., Hawkins G., and Gunson R. (2019). Hepatitis A-2017 an unusual year in Scotland. *J. Clin. Virol*. 115, 1–4.
- Chen N.Y., Liu Z.H., Shie S.S., Chen T.H., and Wu T.S. (2017). Clinical characteristics of acute hepatitis A outbreak in Taiwan, 2015-2016: observations from a tertiary medical center. *BMC Infect. Dis*. 17, 441.
- Chen W.C., Chiang P.H., Liao Y.H., Huang L.C., Hsieh Y.J., Chiu C.M., Lo Y.C., Yang C.H., and Yang J.Y. (2019). Outbreak of hepatitis A virus infection in Taiwan, June 2015 to September 2017. *Euro Surveill*. 24, 1800133.

Cheng C.Y., Wu H.H., Zou H., and Lo Y.C. (2018). Epidemiological characteristics and associated factors of acute hepatitis A outbreak among HIV-coinfected men who have sex with men in Taiwan, June 2015-December 2016. *J. Viral. Hepat.* 25, 1208–15.

Comelli A., Izzo I., Casari S., Spinetti A., Bergamasco A., and Castelli F. (2018). Hepatitis A outbreak in men who have sex with men (MSM) in Brescia (Northern Italy), July 2016-July 2017. *Infez. Med.* 26, 46–51.

Cooksley W.G.E. (2000). What did we learn from the Shanghai hepatitis A epidemic? *Journal of Viral Hepatitis.* 7, 1–3.

Costantino A., Coppola N., Spada E., Bruni R., Taffon S., Equestre M., Marcantonio C., Sagnelli C., Dell’Isola C., Tosone G., et al. (2017). Hepatitis A virus strains circulating during 1997-2015 in Campania, a Southern Italy region with periodic outbreaks. *J. Med. Virol.* 89, 1931–6.

Fares A. (2015). Seasonality of hepatitis: a review update. *J. Family Med. Prim. Care.* 4, 96–100.

Foster M.A., Hofmeister M.G., Kupronis B.A., Lin Y., Xia G.L., Yin S., and Teshale E. (2019). Increase in hepatitis A virus infections — United States, 2013–2018. *Morb. Mortal. Wkly. Rep.* 68, 413–5.

Franco E., Meleleo C., Serino L., Sorbara D., and Zarattil L. (2012). Hepatitis A: epidemiology and prevention in developing countries. *World J. Hepatol.* 4, 68–73.

Freidl G.S., Sonder G.J., Bovée L.P., Friesema I.H., van Rijckevorsel G.G., Ruijs W.L., van Schie F., Siedenburg E.C., Yang J.Y., and Vennema H. (2017). Hepatitis A outbreak among men who have sex with men (MSM) predominantly linked with the EuroPride, the Netherlands, July 2016 to February 2017. *Euro Surveill.* 22, 30468.

Friedman M.S., Blake P.A., Koehler J.E., Hutwagner L.C., and Toomey K.E. (2000). Factors influencing a communitywide campaign to administer hepatitis A vaccine to men who have sex with men. *Am. J. Public Health.* 90, 1942–6.

Gay N.J. (1996). A model of long-term decline in the transmissibility of an infectious disease: implications for the incidence of hepatitis A. *Int J Epidemiol.* 25, 854–61.

Gozlan Y., Bar-Or I., Rakovsky A., Savion M., Amitai Z., Sheffer R., Ceder N., Anis E., Grotto I., Mendelson E., et al. (2017). Ongoing hepatitis A among men who have sex with men (MSM) linked to outbreaks in Europe in Tel Aviv area, Israel, December 2016 - June 2017. *Euro Surveill.* 22, 30575.

de Guimaraens M.A., and Codeço C.T. (2005). Experiments with mathematical models to simulate hepatitis A population dynamics under different levels of endemicity. *Cad. Saude Publica.* 21, 1531–9.

- Heesterbeek J.A.P., and Dietz K. (1996). The concept of R_0 in epidemic theory. *Stat. Neerl.* 50, 89-110.
- Hu M.D., Schenzle D., Deinhardt F., and Scheid R. (1984). Epidemiology of hepatitis A and B in the Shanghai area: prevalence of serum markers. *Am. J. Epidemiol.* 120, 404–13.
- Jacobsen K.H., and Koopman J.S. (2005). The effects of socioeconomic development on worldwide hepatitis A virus seroprevalence patterns. *Int. J. Epidemiol.* 34, 600–9.
- Kermack W.O., and McKendrick A.G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Philos. Trans. Royal Soc. A.* 115, 700-721.
- LaCroix J.M., Snyder L.B., Huedo-Medina T.B., Johnson B.T. (2014). Effectiveness of mass media interventions for HIV prevention, 2006-2013: a meta-analysis. *J. Acquir. Immune Defic. Syndr.* 66 Suppl. 3, S329-40.
- Latash J., Dorsinville M., Del Rosso P., Antwi M., Reddy V., Waechter H., Lawler J., Boss H., Kurpiel P., Bakenson P.B., et al. (2017). Notes from the field: increase in reported hepatitis A infections among men who have sex with men — New York City, January–August 2017. *Morb. Mortal. Wkly. Rep.* 66, 999–1000.
- Lin K.Y., Sun H.Y., Chen Y.H., Lo Y.C., Hsieh S.M., Sheng W.H., Chuang Y.D., Pan S.C., Cheng A., Hung C.C., et al. (2019). Effect of a hepatitis A vaccination campaign during a hepatitis A outbreak in Taiwan, 2015 –2017: a modeling study. *Clin. Infect. Dis.* 70, 1742-1749.
- Ma J. (2020). Estimating epidemic exponential growth rate and basic reproduction number. *Infect. Dis. Model.* 5, 129-141.
- Marosevic D., Belting A., Schönberger K., Carl A., Wenzel J.J., and Brey R. (2019). Hepatitis A outbreak in the general population due to a MSM-associated HAV genotype linked to a food handler, November 2017-February 2018. Germany. *Food Environ. Virol.* 11, 149–56.
- Matheny S.C., and Kingery J.E. (2012). Hepatitis A. *Am. Fam. Physician.* 86, 1027–34, quiz 1010-2.
- Mauro M.V., Greco F., Tenuta R., Apuzzo G., Costantino A., Bruni R., Equestre E., Ciccaglione A.R., Giraldi C., and Mastroianni A. (2019). Hepatitis A outbreak affecting men who have sex with men (MSM) in South Italy. *New Microbiol.* 42, 181–3.
- Ndumbi P., Freidl G.S., Williams C.J., Mårdh O., Varela C., Avellón A., Friesema I., Vennema H., Beebejaun K., Ngui S.L., et al. (2018). Hepatitis A outbreak disproportionately affecting men who have sex with men (MSM) in the European Union and European Economic Area, June 2016 to May 2017. *Euro Surveill.* 23, 1700641.

Nishiura H., and Chowell G. (2009). The Effective Reproduction Number as a Prelude to Statistical Estimation of Time-Dependent Epidemic Trends. *Math. Stat. Estim. Approach Epidemiol.* 103-121.

NIID surveillance data. NIID, Tokyo, 2015-2019. <https://www.niid.go.jp/niid/en/>. Accessed 24 Jul 2019.

Petersen J., Freedman J., Ford L., Gawthrop M., Simons H., Edelstein M., Plunkett J., Balogun K., Mandal S., and Patel D. (2019). Changes to country-specific hepatitis A travel vaccination recommendation for UK travellers in 2017—responding to a vaccine shortage in the national context. *Public Health.* 168, 150–6.

Regan D.G., Wood J.G., Benevent C., Ali H., Smith L.W., Robertson P.W., Ferson M.J., Fairley C.K., Donovan B., and Law M.G. (2016). Estimating the critical immunity threshold for preventing hepatitis A outbreaks in men who have sex with men. *Epidemiol. Infect.* 144, 1528–37.

Rodríguez-Tajes S., Perpiñán E., Caballol B., Lens S., Mariño Z., Costa J., Vilella A., Pérez-Del-Pulgar S., Forns X., and Koutsoudakis G. (2018). Hepatitis A outbreak in Barcelona among men who have sex with men (MSM), January-June 2017: A hospital perspective. *Liver. Int.* 38, 588–93.

Ruscher C., Werber D., Thoullass J., Zimmermann R., Eckardt M., Winter C., and Sagabiel D. (2019). Dating apps and websites as tools to reach anonymous sexual contacts during an outbreak of hepatitis A among men who have sex with men, Berlin, 2017. *Euro Surveill.* 24, 1800460.

Sabrià A., Gregori J., Garcia-Cehic D., Guix S., Pumarola T., Manzanares-Laya S., Caylà J.A., Bosch A., Quer J., and Pintó R.M. (2019). Evidence for positive selection of hepatitis A virus antigenic variants in vaccinated men-having-sex-with men patients: implications for immunization policies. *EBioMedicine.* 39, 348–57.

Takechi A., Hatakeyama S., Kashiya T., and Koibuchi T. (2000). Outbreak of hepatitis A virus infection among men who have sex with men. *Kansenshogaku Zasshi.* 74, 716–9.

Tanaka S., Kishi T., Ishihara A., Watanabe D., Uehira T., Ishida H., Shirasaka T., and Mita E. (2019) Outbreak of hepatitis A linked to European outbreaks among men who have sex with men in Osaka, Japan, from March to July 2018. *Hepatol. Res.* 49, 705–10.

Van Effelterre T.P., Zink T.K., Hoet B.J., Hausdorff W.P., and Rosenthal P. (2006). A mathematical model of hepatitis a transmission in the United States indicates value of universal childhood immunization. *Clin. Infect. Dis.* 43, 158–64.

Van Effelterre T., Marano C., and Jacobsen K.H. (2016). Modeling the hepatitis A epidemiological transition in Thailand. *Vaccine*. 34, 555–62.

Watanabe S., Morimoto N., Miura K., Takaoka Y., Nomoto H., Tsukui M., Isoda N., Ohnishi H., Nagashima S., Takahashi M., et al. (2019). Full-genome characterization of the RIVM-HAV16-090-like hepatitis A virus strains recovered from Japanese men who have sex with men, with sporadic acute hepatitis A. *Hepatol. Res.* 49, 521–30.

Werber D., Michaelis K., Hausner M., Sissolak D., Wenzel J., Bitzegeio J., Belting A., Sagebiel D., and Faber M. (2017). Ongoing outbreaks of hepatitis A among men who have sex with men (MSM), Berlin, November 2016 to January 2017 - linked to other German cities and European countries. *Euro Surveill.* 22, 30457.

WHO position paper on hepatitis A vaccines - June 2012. *Wkly. Epidemiol. Rec.* 87, 261-76.

World Health Organization. Hepatitis A. 2019. <https://www.who.int/news-room/fact-sheets/detail/hepatitis-a>. Accessed 24 Jul 2019.

Zimmermann R., Faber M., Dudareva S., Ingiliz P., Jessen H., Koch J., Marcus U., Michaelis K., Rieck T., Ruscher C., et al. (2021). Hepatitis A outbreak among MSM in Berlin due to low vaccination coverage: Epidemiology, management, and successful interventions. *Int. J. Infect. Dis.* 103, 146-153.

Appendix

The R program used in parameter estimation is shown below.

```
library(ggplot2)

stable_week <- 52

LAMBDA <- 0

hep_data_collect_week <- function(year, week){

  out <- matrix(0, ncol=1, nrow=48)

  for(w in 1:week){

    datapath <- paste("./csv/zenshu_utf-8/", year, "-", formatC(w, width=2, flag="0"), "-
zensu.csv.csv", sep="")

    out <- cbind(out, as.numeric(as.character(read.csv(datapath)[-(1:3), 44])))

  }

  return(out[, -1])

}

hepatitisA_data_2010 <- read.csv("./csv/Syu_03_1_2010.csv", header=F)[c(-1, -2), 2+3*1:52]
hepatitisA_data_2011 <- read.csv("./csv/Syu_03_1_2011.csv", header=F)[c(-1, -2), 2+3*1:52]
hepatitisA_data_2012 <- read.csv("./csv/Syu_03_1_2012.csv", header=F)[c(-1, -2), 2+3*1:52]
hepatitisA_data_2013 <- read.csv("./csv/Syu_03_1_2013.csv", header=F)[c(-1, -2), 2+3*1:52]
hepatitisA_data_2014 <- read.csv("./csv/Syu_03_1_2014.csv", header=F)[c(-1, -2), 2+3*1:52]
hepatitisA_data_2015 <- read.csv("./csv/Syu_03_1_2015.csv", header=F)[c(-1, -2), 2+3*1:52]
hepatitisA_data_2016 <- read.csv("./csv/Syu_03_1_2016.csv", header=F)[c(-1, -2), 2+3*1:52]
```

```

hepatitisA_data_2017 <- read.csv("./csv/Syu_03_1_2017.csv",header=F)[c(-1,-2),2+3*1:52]
hepatitisA_data_2010 <- matrix(as.numeric(as.matrix(hepatitisA_data_2010)),nrow=48)
hepatitisA_data_2011 <- matrix(as.numeric(as.matrix(hepatitisA_data_2011)),nrow=48)
hepatitisA_data_2012 <- matrix(as.numeric(as.matrix(hepatitisA_data_2012)),nrow=48)
hepatitisA_data_2013 <- matrix(as.numeric(as.matrix(hepatitisA_data_2013)),nrow=48)
hepatitisA_data_2014 <- matrix(as.numeric(as.matrix(hepatitisA_data_2014)),nrow=48)
hepatitisA_data_2015 <- matrix(as.numeric(as.matrix(hepatitisA_data_2015)),nrow=48)
hepatitisA_data_2016 <- matrix(as.numeric(as.matrix(hepatitisA_data_2016)),nrow=48)
hepatitisA_data_2017 <- matrix(as.numeric(as.matrix(hepatitisA_data_2017)),nrow=48)
hepatitisA_data_2018 <- hep_data_collect_week(2018,52)
hepatitisA_data_2019 <- hep_data_collect_week(2019,13)

hepatitisA_data_total <-
c(hepatitisA_data_2017[1,],hepatitisA_data_2018[1,],hepatitisA_data_2019[1,])

hepatitisA_data_no_outbreak <-
c(hepatitisA_data_2015[1,],hepatitisA_data_2016[1,],hepatitisA_data_2017[1,])

hep_data <- hepatitisA_data_total

no_outbreak_year <- length(hepatitisA_data_no_outbreak)/length(hepatitisA_data_2010[1,])

moment_to_parm_gamma <- function(x){
  m <- x[1]
  s <- x[2]
  rate <- m / s**2
  shape <- m**2 / s**2
  return(c(shape=shape,rate=rate))
}
41

```

```

}

time_prob_gamma <- function(t_max,param_gamma){
  G <- pgamma(0:t_max,param_gamma[1],param_gamma[2])
  if(any(is.nan(G))){
    browser()
  }
  return(G[2:(t_max+1)] - G[1:t_max])
}

moment_to_parm_lognormal <- function(x){
  m <- x[1]
  s <- x[2]
  meanlog <- log(m**2/(s**2+m**2)**0.5)
  sdlog <- (log(s**2/m**2+1))**0.5
  return(c(meanlog=meanlog,sdlog=sdlog))
}

time_prob_lognormal <- function(t_max,param_lognormal){
  L <- plnorm(0:t_max,param_lognormal[1],param_lognormal[2])
  return(L[2:(t_max+1)] - L[1:t_max])
}

sin_curve <- function(a,n){
  out <- rep(a[1],52)
  if(n > 0){
    for(i in 1:n){
      out <- out + a[2*i] * sin(i * 2 * pi / 52 * 1:52 + a[2*i+1])
    }
  }
}

```

```

    }
  }
  return(out)
}

reproduction_num <- function(R,eps_1,eps_2,t,t_theta,t_1,t_2){
  R_1 <- R
  if(t>t_1){
    R_1 <- R_1 * eps_1
  }
  if(t>t_2){
    R_1 <- R_1 * eps_2
  }
  return(R_1)
}

convolution <- function(u,v){
  return(u %*% rev(v))
}

renewal_proceed <- function(j,w,R_1,eps_1,eps_2,t){
  R <- reproduction_num(R_1,eps_1,eps_2,t,t_theta,t_1,t_2)

  renew <- R * convolution(j,w[1:t])
  return(renew)
}

```

```

renewal_solve <- function(R_1,eps_1,eps_2,w,t_max,t_theta){
  j <- rep(0.0,stable_week + t_theta - 52)
  j[length(j)] <- 1.0
  for(t in (stable_week+t_theta-52):(t_max+stable_week)){
    j_new <- renewal_proceed(j,w,R_1,eps_1,eps_2,t)
    j <- c(j,j_new)
  }
  return(j)
}

```

```

incubation <- function(j,a,f,t){
  incub <- j[1:t] + a[1:t]
  onset <- convolution(incub,f[1:t])
  return(onset)
}

```

```

symptom_onset <- function(j,a,f,t_max,year,start_point){
  E_c <- c()
  env <- c(a[start_point:length(a)],rep(a,year+1))[1:length(j)]
  for(t in (stable_week+1):(t_max+stable_week+1)){
    E_c_new <- incubation(j,env,f,t)
    E_c <- c(E_c,E_c_new)
  }
  return(E_c)
}

```

```

get_likelihood <- function(E_c,hep_data){

  hep_model <- E_c

  hep_model[hep_model == 0] <- 1

  likelihood <- sum(dpois(hep_data,hep_model,log=T))

  return(likelihood)

}

```

```

get_AIC <- function(loglik,parnum){

  return(-2 * loglik + 2 * parnum)

}

```

```

print_result_no_outbreak <- function(a,E_c,t_max,year,start_point){

  height <- 35

  a_plot <- c(a[start_point:length(a)],rep(a,year+1))[1:length(hepatitisA_data_no_outbreak)]

  pdf("plot_no_outbreak.pdf")

  plot(a_plot, xlim=c(0,length(hepatitisA_data_no_outbreak)), ylim=c(0,height), ann=F, type="l", lty=3)

  par(new=T)

  plot(hepatitisA_data_no_outbreak,xlim=c(0,length(hepatitisA_data_no_outbreak)), ylim=c(0,height),
ann=F,type="l",col=3)

  par(new=T)

  plot(E_c, xlim=c(0,length(hepatitisA_data_no_outbreak)), ylim=c(0,height), xlab="week",
ylab="infection", type="l", col=2)

  dev.off()

}

```

```

print_result <- function(j,a,E_c,t_max,year,start_point){

  height <- 35

45

```

```

j_plot <- j[-(1:(stable_week))]

a_plot <- c(a[start_point:length(a)],rep(a,year+1))[1:length(sample_time)]

pdf("plot.pdf")

plot(j_plot[sample_time], xlim=c(0,length(hep_data)), ylim=c(0,height), ann=F, type="l", lty=2)

par(new=T)

plot(a_plot, xlim=c(0,length(hep_data)), ylim=c(0,height), ann=F, type="l", lty=3)

par(new=T)

plot(hep_data,xlim=c(0,length(hep_data)), ylim=c(0,height), ann=F,type="l",col=3)

par(new=T)

plot(E_c, xlim=c(0,length(hep_data)), ylim=c(0,height), xlab="week", ylab="infection", type="l", col=2)

dev.off()

}

nlogL_no_outbreak <- function(a){

  env <- sin_curve(a,depth)

  env <- c(env[start_point:52],env)[1:52]

  E_c <- symptom_onset(rep(0,length(hepatitisA_data_no_outbreak)+stable_week+1),env,
    prob_lognormal_no_outbreak,length(hepatitisA_data_no_outbreak),8,start_point)

  E_c <- E_c[1:length(hepatitisA_data_no_outbreak)]

  likelihood <- get_likelihood(E_c,hepatitisA_data_no_outbreak)

  count <- count + 1

  if(count %% 100 == 0){

    print_result_no_outbreak(env,E_c,t_max,7,start_point)

  }
}

```

```

return(-likelihood)
}

nlogL <- function(parms){
  R_1 <- parms[1]
  eps_1 <- parms[2]
  eps_2 <- parms[3]

  moment_gamma <- c(26.86/7, parms[4]/7)
  parm_gamma <- moment_to_parm_gamma(moment_gamma)
  prob_gamma <- time_prob_gamma(t_max+stable_week,parm_gamma)

  a <- a_init
  env <- rep(a,each=week_interval)
  env <- c(env[start_point:52],env)[1:52]

  j <- renewal_solve(R_1,eps_1,eps_2,prob_gamma,t_max,t_theta)
  E_c <- symptom_onset(j,env,prob_lognormal,t_max,4,start_point)
  E_c <- E_c[sample_time]
  likelihood <- get_likelihood(E_c,hep_data)
  likelihood <- likelihood +
    get_likelihood(rep(env,no_outbreak_year-1),hepatitisA_data_no_outbreak[1:(52 *
(no_outbreak_year - 1))])

  likelihood <- likelihood - LAMBDA * sum((E_c[c(-1,-2)] -

```

$$2 * E_c[c(-1,-length(E_c))] + E_c[c(-length(E_c)+1,-length(E_c))]]**2)$$

```

count <- count + 1

if(count %% 100 == 0){
  out_vector <- c(parms[1],eps_1,eps_2,parms[4],likelihood)
  names(out_vector) <- c("R_1","eps_1","eps_2","sd","likelihood")

  print(out_vector)
  print_result(j,env,E_c,t_max,3,start_point)
}

return(-likelihood)
}

estimation_no_outbreak <- function(a){
  par_N <- a

  opt_L <- optim(par = par_N, fn = nlogL_no_outbreak, gr = NULL, method = "L-BFGS-B", lower = 0, upper
= Inf,

               control = list(parscale = par_N + 1.0e-5), hessian = T)

  return(opt_L)
}

estimation <- function(par_init){
  par_N <- par_init

  opt_L <- optim(par = par_N, fn = nlogL, gr = NULL, method = "L-BFGS-B", lower = 0.1, upper = Inf,

               control = list(parscale = par_N, maxit = 10000), hessian = T)

  return(opt_L)
}

```

```
}
```

```
t_start <- 1
```

```
t_1 <- 9 + 52 - t_start + 1
```

```
t_2 <- 18 + 52 - t_start + 1
```

```
t_1 <- t_1 + stable_week
```

```
t_2 <- t_2 + stable_week
```

```
t_max <- length(hep_data) - 1
```

```
sample_time <- 1:(t_max+1)
```

```
hep_data <- hepatitisA_data_total[t_start:length(hepatitisA_data_total)]
```

```
R_1 <- 1.1
```

```
eps_1 <- 0.5
```

```
eps_2 <- 0.5
```

```
th <- -9
```

```
j_1 <- 1.0
```

```
sd <- 15
```

```
moment_gamma <- c(26.86/7, sd/7)
```

```
parm_gamma <- moment_to_parm_gamma(moment_gamma)
```

```
prob_gamma <- time_prob_gamma(t_max+stable_week,parm_gamma)
```

```
moment_lognormal <- c(31.927342/7,5.259243/7)
```

```
parm_lognormal <- moment_to_parm_lognormal(moment_lognormal)
```

```
prob_lognormal <- time_prob_lognormal(t_max+stable_week+1,parm_lognormal)
```

```
prob_lognormal_no_outbreak <-
```

```
c(prob_lognormal,rep(0,length(hepatitisA_data_no_outbreak)+stable_week+1))[,
```

```
1:(length(hepatitisA_data_no_outbreak)+stable_week+1)]
```

```
week_interval <- 1
```

```
start_point <- 1
```

```
a <- rep(7,ceiling(52 / week_interval))
```

```
setAinit <- function(depth){
```

```
  a <- c(1.0)
```

```
  if(depth > 0){
```

```
    a <- c(a,rep(c(0.1,1.0),depth))
```

```
    a[2] <- 0.3
```

```
  }
```

```
  return(a)
```

```
}
```

```
depth <- 2
```

```
a <- setAinit(depth)
```

```
t_theta <- th + 52 - t_start + 1
```

```
t_theta <- t_theta + stable_week
```

```
count <- 0
```

```
a_init_par <- estimation_no_outbreak(a)$par
```

```
a_init <- sin_curve(a_init_par,depth)
par_init <- c(R_1,eps_1,eps_2,sd)
opt_L <- estimation(par_init)
par_L <- opt_L$par
```