eAppendix 2: Technical modeling details

Simulation and model likelihood

To simulate the model, we fixed values of κ_1 and κ_2 to resolve the non-identifiability of the other parameters; we have verified that all simulations and model outcomes do not depend on these values, only on the values of the identifiable parameter combinations. Initial conditions for the latent L, prodromal P, and jaundiced J compartments of the Southeast Michigan population were chosen so that they were initially in equilibrium with initial condition for the presymptomatic infectious compartment $I_i(0) = I_0$. That is, $L_1(0) = \sigma I_0/\nu$, $P_1(0) = \rho\nu I_0/\mu$, $J_1(0) = \rho\nu I_0/\gamma$. The susceptible and recovered fractions were set to $S_1(0) = 1-L_1(0)-I_1(0)-P_1(0)-J_1(0)$ and $R_1(0) = 0$. The rest of the state was assumed to be disease-free initially, that is $S_2(0) = 1$ and $L_2(0) = I_2(0) = P_2(0) = J_2(0) = 0$.

We modeled the data with a negative binomial distribution to account for overdispersion of the case data $z_i(t_j)$ [4]. Here, the likelihood, which measures the goodness of fit of the model to the data as a function of model parameters θ, is given by

$$\mathcal{L}(\theta) = \prod_{i,j} \frac{1}{z_i(t_j)!} \frac{\Gamma(z_i(t_j) + 1/\alpha_i)}{\Gamma(1/\alpha_i)} \left(\frac{y_i(t_j;\theta)}{y_i(t_j;\theta) + 1/\alpha_i}\right)^{z_i(t_j)} \left(1 + \frac{y_i(t_j;\theta)}{1/\alpha_i}\right)^{-1/\alpha_i}, \tag{1}$$

where α_i is the overdispersion parameter for population i. As $\alpha \to 0$, the distribution becomes Poisson; when $\alpha = 1$, the distribution is geometric; and as $\alpha \to \infty$, the distribution becomes logarithmic.

The 95% likelihood thresholds are given by

$$-\log \mathcal{L}(\hat{\theta}) \pm \frac{1}{2}\chi^2(0.05, m) \tag{2}$$

where $\chi^2(0.05, m)$ is the χ^2 distribution with level of significance 0.05 and number of parameters m=9. There is a region in m-dimensional parameter space that corresponds to model outputs that fit the data within these 95% likelihood thresholds. The profile likelihoods are transects through this confidence region. To find the 95% confidence intervals for the model output trajectories, we simulated the model at points along each profile likelihood, finding the range of model outputs at each time point in this set of simulated trajectories.

Seasonality

We model potential seasonal transmission with logistic functions, which are functions that smoothly transition from 0 to 1; a difference of these functions (as in Eq. (3)) has a smooth, bump-like shape. Parameter δ controls the ratio of peak transmission rate to baseline rate β_0, π_1 controls how steepness of transition, and π_2 and π_3 give the time at which the constituent logistic functions are at their median values. This approach is more flexible than using sinusoidal functions. We then model transmission rates as functions of the calendar week τ,

$$\beta_i(\tau) = \beta_{i,0} \cdot \delta \left(\frac{1}{1 + \exp(-2\pi_1(\tau - \pi_2))} - \frac{1}{1 + \exp(-2\pi_1(\tau - \pi_3))} \right). \tag{3}$$

The existence of seasonal variation in the magnitude of transmission is supported by measures of model fit (difference of 44 in Akaike Information Criterion (AIC) between the best-fit models with and without seasonality). However, a wide range of the seasonality parameters representing the steepness of seasonal transition and the transition timing (π_1, π_2, π_3) were found to be consistent with the data. Thus, we selected parameters representative of the best-fit region for use in the main model, corresponding to increased transmission when the temperature in Southeast Michigan is above 50°F, approximately April through October, with the greatest transition in transmission magnitude occurring over about one month. The main model and the model without seasonal transmission are compared in Figure 1.

Table 1: Seasonality parameters for the Michigan hepatitis A outbreak model

Parameter	Type	Value	Units	Definition
π_1	Fixed	0.34	—	Steepness of seasonal change parameter
π_2	Fixed	22	weeks	Seasonal increase in transmission timing parameter
π_3	Fixed	35	weeks	Seasonal decrease in transmission timing parameter
δ	Estimated	1.6 (1.4–1.7)	—	Magnitude of seasonal variation in transmission

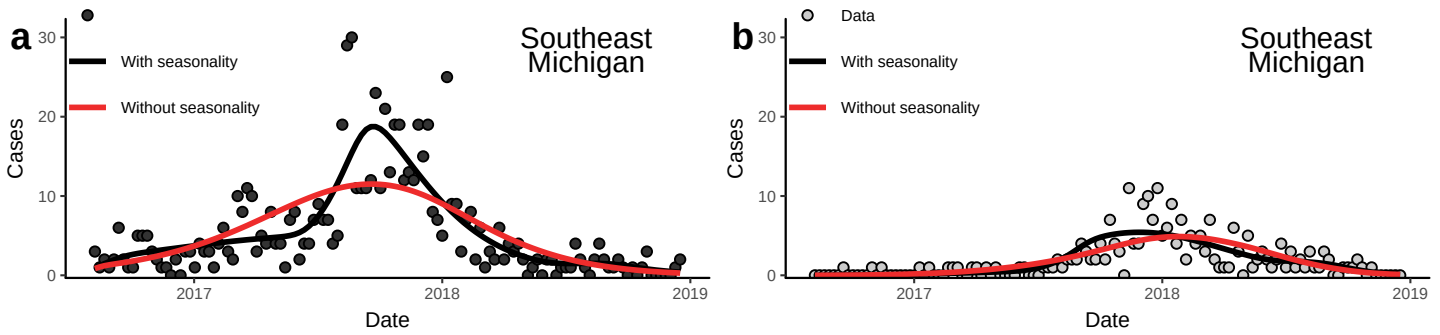

Figure 1: Comparison of model fits with and without seasonal transmission in the Michigan hepatitis A outbreak a) within Southeast Michigan and b) outside of Southeast Michigan.

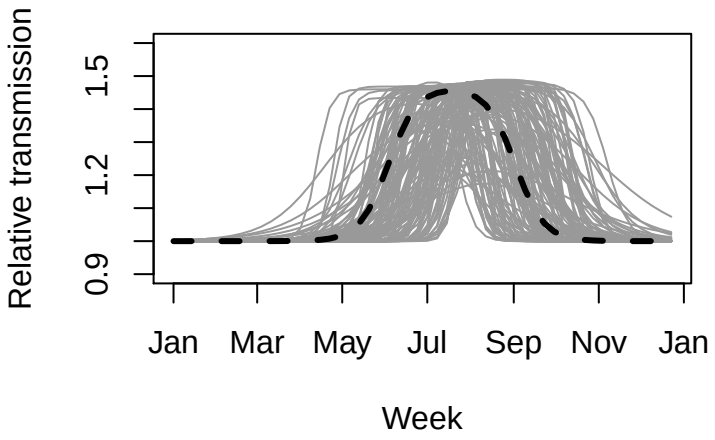

Figure 2: Seasonal variation in transmission of hepatitis A. Each line represents a point in parameter space corresponding to model fits that are within the 95% likelihood-based confidence interval. The dotted line corresponds to the parameters listed in Table 1.

We performed sensitivity analysis on the seasonality parameters by generating test values of π_1, π_2, and π_3 through Latin hypercube sampling with the `lhs` package [1]. We used a classification algorithm and regression tree (CART) method within the `rpart` package [5] to better understand the regions of seasonality parameter space corresponding to model fits within 95% confidence limits. We plot the magnitudes of the seasonal variation in transmission for each of the Latin hypercube samples generated in our sensitivity analysis in Figure 2. The CART algorithm indicated that π_1, the steepness parameter, was largely unimportant for determining whether a seasonality profile fell within the confidence region. The timing parameters were more important. The classification algorithm indicates that good fits have $\pi_2 < 29$ and $\pi_3 > 29$. Moreover, if $\pi_3 > 41$, then $\pi_2 < 26$, but if $\pi_3 < 41$, then $\pi_2 > 16$. Although these values do not have straightforward direct interpretations (due to the functional form of a difference in logistic functions), we can see in Figure 2 that these results indicate a wide range of seasonal patterns are possible, although increased seasonal transmission is unlikely before April and after November.

Previous evidence for seasonality of hepatitis A has been mixed. Incidence in the U.S. is not considered to follow a seasonal pattern [2]. However, globally, there is some evidence for increased transmission in spring and summer months [3]. In the Michigan 2016–19 outbreak, we found that the outbreak trajectory was more consistent with increased transmission (up to around 155% of baseline) during warmer months than with no seasonal pattern to transmission. It may be that person-to-person transmission varies more strongly with temperature than transmission via point sources like food. Moreover, seasonality may only be evident in climates where certain baseline temperatures are not met year-round. Hence, evidence of seasonality may be masked when considering aggregate incidence. It remains to be seen whether there will be evidence of seasonality in other outbreak states.

References

[1] Carnell, R. (2019). *Package 'lhs'*.

[2] Centers for Disease Control and Prevention (2015). *Epidemiology and Prevention of Vaccine-Preventable Diseases, 13th Ed.* Public Health Foundation, Washington, D.C.

[3] Fares, A. (2015). Seasonality of hepatitis: A review update. *Journal of Family Medicine and Primary Care*, 4(1):96.

[4] Lloyd-Smith, J. O. (2007). Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS ONE*, 2(2):1–8.

[5] Thernau, T., Atkinson, B., and Ripley, B. (2019). *Package 'rpart'*.