

## Containing the Spread of Infectious Disease on College Campuses

Mirai Shah<sup>1</sup>, Gabrielle Ferra<sup>2</sup>, Susan Fitzgerald<sup>3</sup>, Paul J. Barreira<sup>4</sup>, Pardis C. Sabeti<sup>5,6,7,8</sup>, Andres Colubri<sup>5,6,7,9\*</sup>

<sup>1</sup> Harvard College, Cambridge, MA, USA

<sup>2</sup> Brown University, Providence, RI, USA

<sup>3</sup> Harvard University Health Services, Cambridge, MA, USA

<sup>4</sup> Harvard Medical School, Boston, MA, USA

<sup>5</sup> Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA, USA

<sup>6</sup> Broad Institute of MIT and Harvard, Cambridge, MA, USA

<sup>7</sup> Howard Hughes Medical Institute, Chevy Chase, MD, USA

<sup>8</sup> Harvard School of Public Health, Boston, MA, USA

<sup>9</sup> University of Massachusetts Medical School, Worcester, MA, USA

\* Corresponding author: [andres.colubri@umassmed.edu](mailto:andres.colubri@umassmed.edu)

### Abstract

College campuses are highly vulnerable to infectious disease outbreaks, and there is a pressing need to develop better strategies to mitigate their size and duration, particularly as educational institutions around the world reopen to in-person instruction in the midst of the COVID-19 pandemic. Towards addressing this need, we applied a stochastic compartmental model to quantify the impact of university-level responses to past mumps outbreaks in college campuses and used it to determine which control interventions are most effective. Mumps is a very relevant disease in such settings, given its airborne mode of transmission, high infectivity, and recurrence of outbreaks despite availability of a vaccine. Our model aims to simultaneously overcome three crucial issues: stochastic variation in small populations, missing or unobserved case data, and changes in disease transmission rates post-intervention. We tested the model and assessed various interventions using data from the 2014 and 2016 mumps outbreaks at Ohio State University and Harvard University, respectively. Our results suggest that in order to decrease infectious disease incidence on their campuses, universities should apply diagnostic protocols that address false negatives from molecular tests, stricter quarantine policies, and effective awareness campaigns among their students and staff. Our model can be applied to data from other outbreaks in college campuses and similar small-population settings.

**Keywords:** Infectious disease, mumps outbreak, college campus, stochastic SEIR model, public health intervention, Harvard University, Ohio State University

# 1 INTRODUCTION

2 The ongoing COVID-19 pandemic has forced school closures around the world (1), and  
3 universities in the United States and elsewhere are designing plans for safe reopening (2, 3). This  
4 is a challenging task, as college campuses provide ideal breeding grounds for infectious disease.  
5 Students live in close quarters, pack into lecture halls, share food and drinks in the dining areas,  
6 and engage in intimate contact. Outbreaks in these settings can spread very quickly. Indeed, a  
7 meningitis outbreak took place at Princeton University in March 2014, eventually claiming the life  
8 of one student. The Centers for Disease Control and Prevention (CDC) reported the attack rate of  
9 the disease on Princeton’s campus to be 134 per 100,000 students – 1,400 times greater than the  
10 national average (4).

11 A recent string of outbreaks on college campuses involves mumps, once a common  
12 childhood viral disease. After introduction of the measles-mumps-rubella (MMR) vaccine in 1977  
13 and the two-dose MMR vaccination program in 1989, the number of mumps cases in the US  
14 plummeted by 2005. But, despite a vaccinated population, there has been a recent resurgence of  
15 mumps, with a steep jump from 229 cases in 2012 to 5833 cases in 2016 (5). Although a typically  
16 mild disease in children, up to 10% of mumps infections acquired after puberty can cause severe  
17 complications, including orchitis, meningitis, and deafness. Furthermore, a majority of recent  
18 mumps cases have occurred in young adults who had received the recommended two MMR doses.  
19 This suggests that vaccine-derived immunity wanes over time, unlike natural immunity –  
20 protection acquired from contracting the disease – which is permanent. Lewnard and Grad estimate  
21 that 33.8% of young adults (ages 20 to 24) were susceptible to mumps in 1990, in contrast to the  
22 52.8% susceptible in 2006, as vaccinations have replaced contraction as the source of immunity  
23 (6). The temporary immunity from vaccines strengthens the argument for strict containment as a

24 critical line of defense amidst an outbreak. In the case of COVID-19, even with the availability of  
25 several vaccines (7), the challenges associated with their wide and quick distribution (8), the  
26 substantial asymptomatic and pre-symptomatic transmission of the disease (9), and the possibility  
27 of new viral strains with higher transmissibility (10) provide further support for such approaches.

28         The spread of mumps at Harvard University in 2016, and extensive public health measures  
29 and documentation, presents a rare opportunity to closely examine an outbreak on a college  
30 campus. Between January 1 and August 31, 2016, 210 confirmed mumps cases were identified in  
31 the Greater Boston area, with most detected at Harvard University. Mumps is a highly contagious  
32 disease with the potential to travel quickly and pervasively on a crowded college campus. Some  
33 of the most notable mumps outbreaks on college campuses occurred in Iowa (11), Indiana (12),  
34 and Ohio (13). But, whereas mumps spread rapidly at Ohio State University (OSU) in 2014 and  
35 the University of Iowa in 2006 and 2016, Harvard employed a number of interventions that may  
36 have helped mitigate spread of the disease and contain it over just a few months (14). The  
37 possibility of distinct viral strains resulting in different outbreak dynamics between schools can be  
38 safely dismissed, as it was shown by application of genetic epidemiology methods (15) that all  
39 mumps outbreaks in the US since at least 2006 have been likely caused by the same mumps lineage,  
40 mumps virus genotype G.

41         The successful containment at Harvard motivates us to explore varied intervention  
42 strategies, given the relative costs of prevention. Even if the use of a booster MMR vaccination is  
43 proven theoretically to reduce infection and thus potentially prevent outbreaks (6, 11), it is unlikely  
44 that universities with limited resources will proactively invest in a third dose. A rough cost analysis  
45 conducted by Harvard University Health Services (HUHS) showed that, while the total mumps  
46 care expenses for Harvard was approximately \$75,000, the cost of providing a third MMR dose to

47 every member of the Harvard community (at \$83 per vaccination) was \$1.7 million (16). Therefore,  
48 at least in the short term, a third MMR dose cannot be the only answer to handling mumps  
49 outbreaks; we must consider more immediate solutions and interventions.

50 In order to understand the effectiveness of interventions aimed at containing an outbreak  
51 on a college campus, we constructed an epidemiological model to simulate the dynamics of mumps  
52 on such a population and quantify the impact of various interventions. Most epidemiological  
53 models have at least one of three flaws: they cannot handle random fluctuations in a small  
54 population, require complete data without unobserved or missing cases, or do not accommodate  
55 time-varying infection or recovery rates as a result of dynamically changing interventions. The  
56 modified stochastic susceptible-exposed-infectious-recovered (SEIR) model presented in this  
57 paper addresses these three issues. We developed this model within the framework of a Partially  
58 Observed Markov Process (POMP), which has been applied to introduce structural stochasticity  
59 into epidemic models (17). We fit model parameters on case data for Harvard's 2016 mumps  
60 outbreak provided by the Massachusetts Department of Public Health (MDPH). We compared it  
61 to data from OSU, one of the few universities that had extensive publicly available data through  
62 the CDC.

63 In applying our model, we found that each of the interventions employed by HUHS -- email  
64 awareness campaigns, more aggressive diagnoses where clinical symptoms alone were enough to  
65 result in quarantine, and strict isolation of suspected cases -- were crucial in reducing the size and  
66 duration of the outbreak. In particular, Harvard's policies drastically increased the reporting rate  
67 of infection and shortened the time a person remains infectious in a susceptible population, relative  
68 to the baseline. As a result, one mumps case at Harvard infected less than two susceptible  
69 individuals on average, and much less once aggressive diagnosis was in place, compared to cases

70 at non-residential schools like OSU, in which one mumps case infected an average of six  
71 susceptible individuals. However, the OSU data suggests that self-isolation could be effective, if  
72 adopted rigorously by students. The conclusions from this paper could guide future responses to  
73 infectious disease outbreaks on college campuses. Without effective measures in place, highly  
74 transmissible diseases like mumps, meningitis, and now COVID-19, spread in these environments  
75 at much faster rates than in the overall population and can lead to serious health complications.  
76 Simple interventions that ensure most cases are detected, treated, and separated from susceptible  
77 individuals make a significant difference.

78

## 79 **2. MATERIALS AND METHODS**

80

### 81 **2.1 Harvard mumps outbreak**

#### 82 **2.1.1 Data**

83 The mumps outbreak at Harvard began in February 2016, when six students reported onset of  
84 parotitis to HUHS. For the next three months, the number of cases continued to rise, until finally  
85 plateauing in late May and early June. There were two waves of the outbreak – one occurring in  
86 the month of March and a larger one occurring in mid-April – totaling 189 confirmed and probable  
87 cases (Figure 1). Confirmed cases are those with a positive laboratory test for mumps virus.  
88 Probable cases are those who either tested positive for the anti-mumps IgM antibody or had an  
89 epidemiologic linkage to another probable or confirmed case (18, 19). The majority of these cases  
90 received the recommended two doses of MMR (20).

91 We use data provided by MDPH, which documented every mumps case between 2015 and  
92 2017 at schools across Massachusetts (21). This data includes demographics of the patient (gender,  
93 age, county, and institution), symptoms and vaccination status, date they reported their symptoms  
94 and the date of symptom onset, and lag time between the date of symptom onset and admission to  
95 a medical clinic.

96

#### 97 **2.1.2 Interventions**

98 Harvard University employed three main interventions: (i) an email awareness campaign, (ii) more  
99 aggressive diagnoses, and (iii) strict isolation of infectious persons.

100 First, between February and May 2016, HUHS sent six different emails to Harvard students,  
101 employees, and colleagues with information on the gravity of the outbreak, recommendations on  
102 how to prevent transmission, and instructions on how to identify mumps. This raised awareness

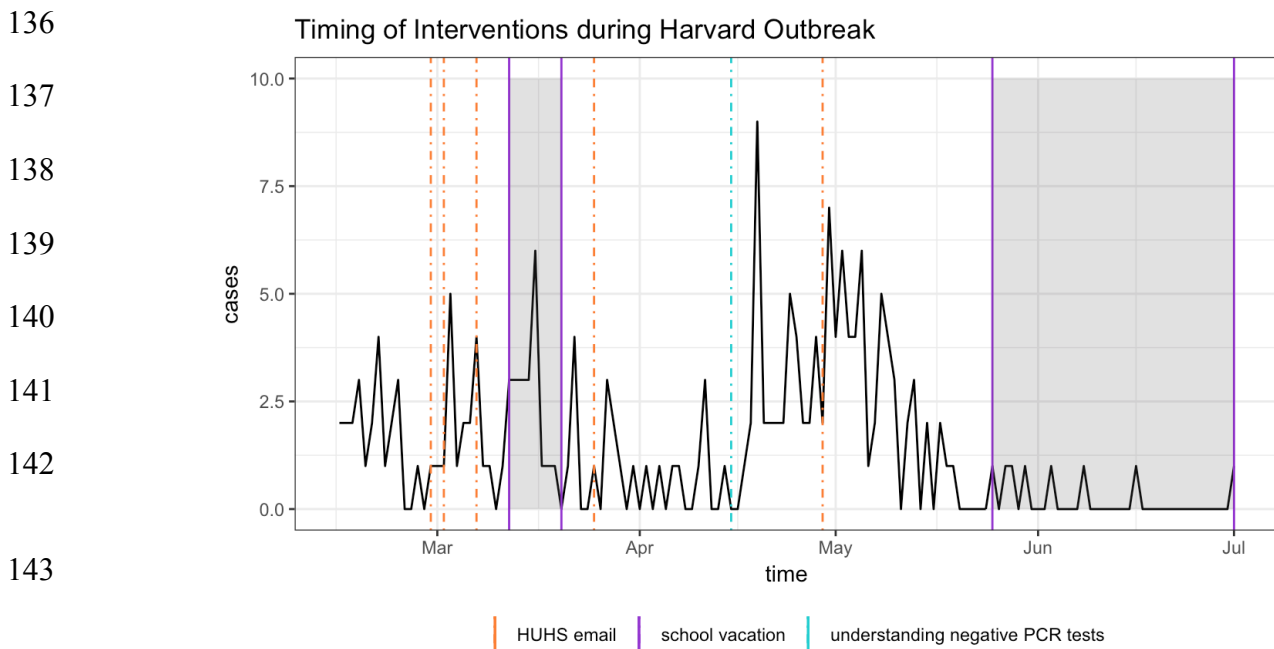
103 throughout the campus. Particularly at the peak of the outbreak, roommates, resident deans, and  
104 athletic coaches all played essential roles in reporting potential cases of mumps, so that few cases  
105 likely went undetected and untreated by HUHS (18, 19).

106 Second, Harvard acted vigorously to treat and isolate anyone suspected of mumps  
107 throughout the outbreak. Initially, due to the disease's non-specific symptoms and less extreme  
108 manifestation in vaccinated people, HUHS used positive mumps PCR tests as a necessary ground  
109 for diagnosis. Later, on recommendation from the MDPH, HUHS stopped automatically ruling out  
110 those with negative PCR results, given that false negatives were quite frequent in vaccinated  
111 individuals and that some individuals reported their infection to the clinic belatedly. In outbreaks  
112 among two-dose vaccine recipients, mumps virus was only detected in samples from  
113 approximately 30-35% of case patients if the samples were collected within the first three days  
114 following onset of parotitis (22). Anyone who entered HUHS displaying clinical symptoms of  
115 mumps was now deemed infected and infectious. This change in the diagnosis protocol took place  
116 on April 15 2014, day 61 of the outbreak (19).

117 Third and perhaps most notably, Harvard isolated most confirmed or probable cases of  
118 mumps. While many universities simply suggest self-isolation in one's room or dormitory (which  
119 leaves roommates and friends highly susceptible to the disease), Harvard removed anyone with  
120 clinical symptoms of mumps from the population. Of the 230 total cases at Harvard between  
121 February 2016 and November 2017, 96 were isolated in alternate housing on campus, while 110  
122 were isolated off-site. Although a person remains infectious with mumps for five days, Harvard  
123 isolated patients for six days for additional measure (18).

124 Harvard also used a variety of smaller techniques to contain the disease. For instance, water  
125 fountains with a weak upward flow were repaired in late March when it became apparent that

126 students were directly touching the fountain with their water bottles or mouths (19). In this study,  
127 we only considered the first three larger-scale interventions in our models. Figure 1 shows a  
128 timeline of the interventions as well as periods when the population was fluctuating (such as during  
129 spring and summer break). Around two weeks after HUHS improved its criteria for diagnosis in  
130 mid-April, there was a steep decline in the number of new cases. These interventions were possible  
131 thanks to the ample resources that Harvard has at its disposal, which may not be available at other  
132 universities. Nevertheless, this situation makes Harvard an ideal testing ground for interventions  
133 that could not be deployed elsewhere, at least without solid proof of their efficacy. Thus, we  
134 quantify the effects of the three main interventions (awareness campaign, aggressive diagnoses,  
135 and strict isolation of suspected cases) further in the modeling section of this paper.



**Figure 1:** The daily number of new mumps cases (probable or confirmed) at Harvard and the timeline of school vacations and control interventions employed by HUHS between February and June 2016. Both probable and confirmed cases display clinical symptoms of mumps, but only confirmed cases have a positive PCR result. HUHS sent multiple emails over the course of the outbreak, raising awareness about the spread of mumps. Additionally, in mid-April, HUHS began more carefully diagnosing mumps, rather than automatically ruling out those with negative PCR tests. The isolation policy is not shown because it occurred continuously throughout the entire outbreak.



## 149 **2.2 Ohio State University mumps outbreak**

### 150 **2.2.1 Data on the outbreak**

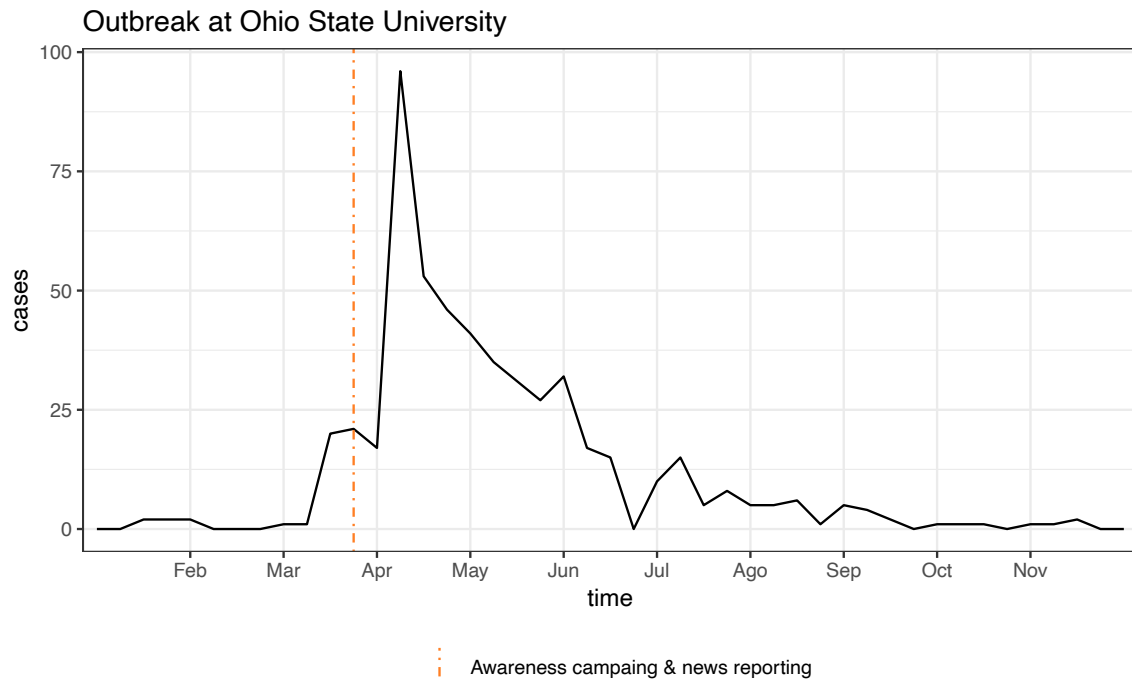
151 In 2014, a large outbreak of mumps occurred in central Ohio, with the majority of cases linked to  
152 OSU in Columbus. The outbreak began in February 2014 and peaked in early April with 96 cases  
153 in one week. By summer and early fall, the number of cases had dramatically dropped and  
154 stabilized (13). We therefore restrict our analysis of the outbreak to the time between Week 1 and  
155 Week 40 of 2014, in which there were a total of 528 cases (Figure 2). We obtained this data from  
156 CDC's *Morbidity and Mortality Weekly Report* (23). One drawback of the data is that the cases  
157 are reported weekly, making our analysis and parameter estimations less precise. Furthermore, we  
158 cannot guarantee that all the cases in this dataset are linked to the university itself, but we know  
159 from news reports that most cases in Ohio occurred on campus during the first half of 2014 (13).  
160 The proximity in time to the Harvard outbreak and the differences in response detailed below make  
161 this a good dataset to compare to.

162

### 163 **2.2.2 Characteristics of the response**

164 We were unable to acquire data directly from OSU, and thus the exact timeline and range of  
165 interventions administered over this period are not known. We learned through online searches  
166 that advisories were published by the university, notifying students of the issue and how to prevent  
167 its spread. One notice published by OSU's medical center reads: "Stay at home for five days after  
168 symptoms (salivary gland swelling) begins (required by Ohio law OAC 3701-3-13, (P)); avoid  
169 school, work, social gatherings, and other public settings" (24). These advisories were distributed  
170 since March 2014 (25), and local news outlets also started reporting the outbreak earlier in the

171 month (26). It appears, however, that like most affected universities, OSU did not formally isolate  
172 infectious persons.



173

174

175

176

**Figure 2:** Number of weekly mumps cases in Ohio (particularly Ohio State University) between January and November 2014. There were 528 cases during this time period, with most occurring between March and July. The dotted line in the last week of March indicates the intervention consisting in awareness campaign by OSU, as well as local and national news reports about the outbreak.

### 177 **2.3 Epidemiological POMP model**

178 The epidemiology of mumps can be captured by a Susceptible-Exposed-Infected-Removed (SEIR)

179 compartmental model: after exposure, individuals go through a latent non-infectious period,

180 followed by an infectious phase (27). Infectious individuals are removed from the transmission

181 process either by recovery or isolation, after which they become immune. Compartmental models

182 simplify the mathematical modeling of infectious diseases; however, they assume access to fully

183 observed disease data. In reality, not all mumps cases are reported, and latent mumps carriers

184 exhibit no symptoms at all. In order to address this issue, our approach integrates a standard SEIR

185 model with a Partially Observed Markov Process (POMP) model (28). This allows us to combine  
186 the simplicity of compartmental models with a probabilistic framework for the underlying  
187 dynamics and the observed data. POMP models require the specification of a process model that  
188 describes stochastic transitions between the (unobserved) states of the system (in this case, the  
189 SEIR compartments), and a measurement model where the distribution of observed data (e.g.:  
190 confirmed cases) is expressed as a function of the unobserved states. The stochasticity introduced  
191 in the SEIR dynamics makes our model better suited to describe small populations, such as college  
192 campuses, where random fluctuations can be significant in relation to the size of the population.  
193 We describe the process and measurement models below.

194

### 195 **2.3.1 Process model**

196 The process model, defined as a stochastic SEIR model, provides the change in true incidence of  
197 mumps at every time point. We add parameters that induce random fluctuations into the population  
198 and change the compartments' rates of transfer in response to interventions. We do this by using  
199 probabilistic densities for the transition of state variables. Moreover, although disease dynamics  
200 are technically a continuous Markov process, this is computationally complex and inefficient to  
201 model, and so we make discretized approximations by updating the state variables after a time step,  
202  $\delta$ . Due to the varying granularity of the observed data (daily and weekly), we used two different  
203 time steps:  $\delta_H = 2.4 \text{ hours}$  for Harvard and  $\delta_O = 12 \text{ hours}$  for OSU. The system of discretized  
204 equations is shown in Equation 1, where  $B(t)$  is the number of susceptible individuals who  
205 become exposed to mumps,  $C(t)$  is the number of newly infectious cases, and  $D(t)$  is the number  
206 of cases that are removed from the population:

$$\begin{aligned} 207 \quad & S(t + \delta) = S(t) - B(t) \\ 208 \quad & \\ 209 \quad & E(t + \delta) = E(t) + B(t) - C(t) \\ 210 \quad & \\ 211 \quad & I(t + \delta) = I(t) + C(t) - D(t) \quad (1) \\ 212 \quad & \\ 213 \quad & R(t + \delta) = R(t) + D(t) \\ 214 \quad & \\ 215 \quad & S(t) + E(t) + I(t) + R(t) = N \\ 216 \quad & \end{aligned}$$

217 Equation 1 describes how the sizes of the four compartments (susceptible, exposed,  
218 infectious, and removed) change between  $(t, t + \delta)$ . The model further assumes that the  
219 population size  $N$  remains constant at every time point. We added inherent randomness to our  
220 model by setting  $B(t)$ ,  $C(t)$ , and  $D(t)$  as binomials. If we assume that the length of time an  
221 individual spends in a compartment is exponentially distributed with some compartment-specific  
222 rate  $x(t)$ , then the probability of remaining in that compartment for an additional day is  
223  $\exp(-x(t))$  and the probability of leaving that compartment is  $1 - \exp(-x(t))$ :

$$\begin{aligned} 224 \quad & B(t) \sim \text{Bin}(S(t), 1 - \exp(-\lambda(t))), \text{ where } \lambda(t) = \beta(t) \frac{I(t)}{N} \\ 225 \quad & \\ 226 \quad & C(t) \sim \text{Bin}(E(t), 1 - \exp(-\sigma)) \quad (2) \\ 227 \quad & \\ 228 \quad & D(t) \sim \text{Bin}(I(t), 1 - \exp(-\gamma(t))) \end{aligned}$$

229 The force of infection,  $\lambda(t)$ , is the transition rate between the susceptible and exposed  
230 classes at time  $t$ , and can be expressed as  $\beta(t) \frac{I(t)}{N}$ , where  $\beta(t)$  represents the transmission rate of  
231 the disease. The removal rate between the infectious and removed compartments at time  $t$  is given  
232 by  $\gamma(t)$ , and transition rate between the exposed and infectious classes is  $\sigma$ . Therefore,  $\gamma(t)^{-1}$   
233 represents the mean length of time a person is infectious before being removed from the population  
234 (either because of intervention efforts or natural recovery), while  $\sigma^{-1}$  represents the mean length  
235 of time a person stays in the latent stage. With this notation, we are implicitly assuming that the  
236 transmission and removal rates could change over time due to interventions or changes in behavior,

237 while the duration of the latent stage is constant and determined by the physiopathology of the  
238 disease. We will justify these assumptions for Harvard and OSU next, as well as provide explicit  
239 formulas for  $\beta(t)$  and  $\gamma(t)$ .

240 Leaving aside the unlikely possibility of change in pathogen's infectivity, the transmission  
241 rate  $\beta(t)$  essentially depends on the frequency of exposure events. In the case of Harvard, its  
242 nature as a residential campus would lead to significant decreases in student population, and  
243 therefore exposures, during school vacations. Exposure at OSU, a non-residential campus, is  
244 arguably less affected by vacation breaks. Another potential cause for reduction in exposures is  
245 awareness campaigns resulting in the adoption of preventive behaviors by students. Both Harvard  
246 and OSU adopted such campaigns, in the former, implemented as emails regularly sent out by  
247 HUHS recommending personal hygiene and testing in case of symptoms compatible with mumps;  
248 in the latter, in the form of advisories posted around campus and online, advising self-isolation to  
249 those students who presented symptoms. Furthermore, due to the scale of the mumps outbreak in  
250 Ohio, it received local and national news coverage, particularly in connection with OSU.  
251 Anecdotal evidence (i.e.: conversation with students) and, most importantly, the fact that HUHS  
252 emails were throughout the outbreak, make us conclude that emails were not particularly effective.  
253 On the other hand, news coverage in the case of OSU could have led to additional awareness by  
254 students and encouraged some to self-isolate. We argue that self-isolation results in lowering of  
255 transmission rate, not shortening of the removal time, because it is not perfect quarantine and  
256 people can still interact and become exposed, albeit at a lower frequency. Based on these known  
257 facts and our interpretation of them, we propose the following transmission rate  $\beta_H(t)$  for the  
258 Harvard model:

$$\begin{aligned} 259 \quad \beta_H(t) &= p\beta_H, \quad t_0 \leq t \leq t_1 \text{ or } t \geq t_2 \\ 260 \quad &= \beta_H \text{ otherwise} \end{aligned} \quad (3)$$

261 Here,  $t_0$  and  $t_1$  represent the starting and ending dates for the spring break (March 12-20 2016),  
262 and  $t_2$  the beginning of the summer recess (May 26 2016). The constant  $\beta_H$  is the baseline  
263 transmission rate during normal class term, and the parameter  $p$  is a number between 0 and 1 that  
264 accounts for the reduction of student population on campus during the school vacation. In the case  
265 of OSU, we propose:

$$\begin{aligned} 266 \quad \beta_o(t) &= w\beta_o, \quad t \geq \zeta \\ 267 \quad &= \beta_o \text{ otherwise} \end{aligned} \quad (4)$$

268  
269 In this equation,  $\beta_o$  the baseline transmission rate,  $w$  is a constant lower than 1, and  $\zeta$  the time  
270 when students began to self- quarantine. Based on publication of public health advisories and  
271 local news, we set this time as the last week of March 2014 (week 12). Since Harvard's  
272 quarantine was in effect through the entirety of the outbreak, we did not incorporate a similar  $w$   
273 coefficient to the corresponding  $\beta_H(t)$  equation for Harvard.

274 The removal rate  $\gamma(t)$  can also be affected by interventions and personal behaviors. We  
275 know that HUHS diagnosis protocol changed on day 61 of the outbreak at Harvard, resulting in a  
276 shorter average removal time since clinical presentation of symptoms alone was enough to result  
277 in strict isolation of suspected cases. Thus, we propose the following  $\gamma_H(t)$  for Harvard:

$$\begin{aligned} 278 \quad \gamma_H(t) &= q\gamma_H, \quad t \geq \tau \\ 279 \quad &= \gamma_H, \quad t < \tau \end{aligned} \quad (5)$$

280 Here,  $q$  is a constant greater than 1 and  $\tau$  is the date when the new criteria was implemented (April  
281 15, 2014). The constant  $\gamma_H$  is the baseline removal rate reflecting the impact of the original  
282 diagnosis protocol. In the OSU model, on the other hand, we assume a constant recovery rate  $\gamma$   
283 equal to the population average for mumps, since infected individuals self-isolate at home. This  
284 would not result in a strict quarantine but in a reduced contact rate with susceptible individuals,  
285 which is already modeled by a lower transmission rate in equation (4).

286 Finally, it is necessary to estimate the basic reproduction number,  $R_0$ , which equals the  
287 expected number of secondary cases produced by an infectious person in a completely susceptible  
288 population (27).  $R_0$  measures the initial growth rate of an outbreak and so, if it is less than 1, then  
289 the infection will die out and there will be no epidemic. For our stochastic SEIR model, this  
290 constant can be expressed as  $R_0 = \frac{\beta}{\gamma}$  (29). Meanwhile, the time-dependent effective reproduction  
291 number is defined as  $R_E(t) = \frac{\beta(t)}{\gamma(t)} * \frac{S(t)}{N}$ , but because  $S(t) \approx N$ , we can simplify this expression  
292 to  $R_E(t) \approx \frac{\beta(t)}{\gamma(t)}$ . Both the basic and effective reproduction numbers allow us to understand the  
293 strength of an outbreak.

294

### 295 **2.3.2 Measurement Model**

296 Although it is impossible to directly record the number of people that are susceptible, exposed,  
297 infectious, and removed directly, the MDPH and CDC data tells us the number of observed cases  
298 per day. The mean number of observed cases per day is the true number of cases multiplied by the  
299 reporting rate  $\rho$  ( $\rho < 1$ ). However, rather than simply denoting the observed number of cases as a  
300 binomial distribution, we account for greater variability in the measurements than a binomial  
301 distribution expects, since college populations are “small” (compared to cities and larger  
302 administrative units) and more affected by random fluctuations (30). Thus, the number of observed  
303 cases,  $y_t$ , given the number of true cases,  $C(t)$ , can be best modelled by an overdispersed binomial  
304 distribution defined as a discretized Normal random variable:

$$305 \quad y_t | C(t) \sim Normal(\rho C(t), \rho(1 - \rho)C(t) + (\psi \rho C(t))^2) \quad (6)$$

306

307

308 The parameter  $\psi$  handles the increased variability in a small population. If  $\psi = 0$ , the  
309 variance in our measurement model simplifies to the variance for a binomial distribution.

310

### 311 **2.3.3 Final POMP Model**

312 The process and measurement models define our final POMP model. For each time point, the  
313 process model generates the number of new cases based on binomially distributed counts. The  
314 measurement model then estimates the observed number of cases based on the true number of  
315 cases and reporting rate. The free parameters in our POMP models for Harvard and OSU that need  
316 to be estimated from the data are the following: (i)  $\beta_H$  and  $\beta_O$ , baseline transmission rates, (ii)  $p$   
317 and  $w$ , decrease in transmission rate at Harvard and OSU due to vacation and self-isolation,  
318 respectively, (iii)  $\gamma_H$  baseline removal rate at Harvard (iv)  $q$ , increase in removal rate due to the  
319 updated HUHS diagnosis protocol, (v)  $\rho_H$  and  $\rho_O$ , case reporting rates, (vi)  $\psi_H$  and  $\psi_O$  ,  
320 overdispersion coefficient representing additional variability in the populations.

321

## 322 **2.3 Fixed parameters**

323 In addition to the free parameters to be estimated from the observed case data, our models also  
324 include a number of fixed parameters, shown in Table 1, whose values can be inferred directly  
325 from previous knowledge or available information. As mentioned earlier, we chose  $\tau = 61$  days  
326 and  $\zeta = 12$  weeks because those points in time at Harvard and OSU correspond to the introduction  
327 of the interventions that we hypothesized to be impactful in the dynamics of the respective  
328 outbreaks. Dates  $t_0$ ,  $t_1$ , and  $t_2$  for the spring and summer vacations at Harvard are available online  
329 (31). We set the rate between the exposed and infectious classes and the recovery rate to  $\sigma = \frac{1}{17}$   
330 and  $\gamma = \frac{1}{5}$ , respectively, since the average latent period and recovery time for mumps are known  
331 to be  $\sigma^{-1} = 17$  days and  $\gamma^{-1} = 5$  days (6). Finally, we set the effective population size at Harvard  
332  $N_H = 20,000 \times 0.53 = 10,600$  people based on records of Harvard's enrollment and



333 employment, and Grad and Lewnard estimation of susceptibility to mumps among college-age  
 334 adults due to immunity waning (6). Similarly, we use an effective population for OSU given by  
 335  $N_O = 60,000 \times 0.53 = 31,800$ , leveraging the total enrollment for the 2013-2014 academic year  
 336 reported in OSU’s statistics website (32).

337

<i>Symbol</i>	<i>Description</i>	<i>Value</i>	<i>Units</i>	<i>Source</i>
$\tau$	Date of intervention at Harvard	61	day	Harvard records on interventions (19)
$t_0, t_1, t_2$	Vacation dates at Harvard, 2015-2016 academic year	26, 34, 100	day	Harvard archived academic calendar (31)
$\zeta$	Date of intervention at OSU	12	week	
$\sigma^{-1}$	Duration of mumps latent period	17	day	Lewnard and Grad (6)
$\gamma^{-1}$	Duration of mumps recovery period	5	day	Lewnard and Grad (6)
$N_H$	Effective population at Harvard	10,600	—	Harvard records on population size (20) and mumps susceptibility among college-aged individuals (6)
$N_O$	Effective population at OSU	31,800	—	OSU’s statistical summary (32) and mumps susceptibility among college-aged individuals (6)

338 **Table 1:** List of fixed parameters used in mumps transmission model for Harvard and OSU

339

## 340 **2.4 Maximum likelihood estimation of free parameters**

341 In order to obtain estimates of the free parameters in our models, we pick the parameter values that  
 342 maximize the log likelihood of the observed data given each model. Within the POMP framework,  
 343 we can perform fast maximum likelihood estimation (MLE) via Sequential Monte Carlo (SMC)  
 344 techniques (28). SMC allows us to calculate the likelihood of the data more efficiently by applying  
 345 the Markov property to generate paths in parameter space that sample the likelihood surface. We  
 346 performed 100 searches from random parameter guesses, each converging to a unique value, and  
 347 we then took the maximum over the 100 runs the final point estimates. We did this using the pomp  
 348 package version 2.8 (33) for the R statistical software version 3.6.1 (34). In order to calculate the  
 349 confidence intervals for each parameter, we selected the top quartile from the set of parameters

350 values obtained in the SMC runs, and applied the adjusted bootstrap percentile (BCa) method (35)  
351 with 10,000 bootstrap replicates using the function `boot.ci` in version 1.3.20 of package `boot` for R  
352 (36).

353

## 354 **2.5 Intervention analysis**

355 Finally, we performed an analysis of the parameters  $q$  and  $w$ , which respectively quantify the effect  
356 of what we consider to be the defining intervention at Harvard (aggressive diagnosis) occurring  
357 around day 61 of the outbreak, and the self-isolation awareness campaign at OSU during March  
358 2014. This could allow us to understand to what extent these interventions made a difference on  
359 the trajectory of the outbreak. First, we compared the scenario with the interventions versus a  
360 scenario without the interventions. Controlling for all other parameters, we run two sets of  
361 simulations at the MLEs, with 200 simulations each. The first set of simulations fixed  $q$  and  $w$  at  
362 the value obtained from MLE, while the second set of simulations set  $q$  and  $w$  to 1, assuming that  
363 no interventions occurred around day 61 at Harvard and by week 12 at OSU. We then compared  
364 the cumulative number of cases over time for these two sets of simulations, generating a 95%  
365 percentile range from all the simulations in each set. Second, we used this method to determine if  
366 administering the interventions earlier could have lowered the number of cases. For Harvard, we  
367 let the day of the intervention take on values between 1 and 60. Subsequently, we ran simulations  
368 for each of these 60 cases, pulled the final outbreak size from the median simulation, and calculated  
369 the reduction in outbreak size. We applied the same procedure for OSU, in this case varying the  
370 day of intervention between 1 and 11 and calculating the corresponding final outbreak sizes.

### 371 3. RESULTS

#### 372 3.1 Optimal Parameters of Harvard and OSU Outbreaks

373 The MLEs of the parameters provide insight into the key characteristics of Harvard’s and OSU’s  
 374 outbreak. In general, we observe very good agreement between the observed cases and the  
 375 simulated outbreaks using the optimal parameters. The effective reproduction number also reflects  
 376 the effects of the interventions at Harvard and OSU in way that’s consistent with our initial  
 377 modeling assumptions. The bootstrap sampling method results in narrow 95% CIs.

378

##### 379 3.1.1 Maximum Likelihood Estimates for Harvard

380 The results are shown in Table 2. Notably, the baseline removal rate  $\gamma_H$  is quite high, indicating  
 381 that the initial diagnosis protocol was quite effective at identifying and removing infected students  
 382 from the population, but it was further increased after day 61. The reporting rate  $\rho_H$  is also  
 383 remarkably high, which suggests that HUHS was able to identify most of the cases circulating at  
 384 Harvard.

385

<i>Symbol</i>	<i>Description</i>	<i>Point estimate</i>	<i>95% CI</i>	<i>Units</i>	<i>Source</i>
$\beta_H$	Baseline transmission rate	1.39	(1.29, 1.42)	day <sup>-1</sup>	MLE
$\gamma_H$	Baseline removal rate	0.85	(0.81, 0.88)	day <sup>-1</sup>	MLE
$p$	Decrease in infection due to vacation	0.11	(0.09, 0.15)		MLE
$q$	Increase in removal rate	2.8	(2.25, 2.52)	—	MLE
$\rho_H$	Proportion of infections reported	0.97	(0.92, 0.95)	—	MLE
$\psi_H$	Overdispersion parameter	0.54	(0.49, 0.53)	—	MLE
$R_E(t)$	Effective reproduction number	1.63 normal term 0.18 during vacation 0.58 after intervention	—	—	Calculated as $\frac{\beta(t)}{\gamma(t)}$

386

**Table 2:** List of parameters in the Harvard model that were obtained by MLE or calculated using the estimated parameters.

387

388 We ran stochastic simulations of Harvard’s outbreak using the parameter values from Table  
 389 2. Figure 3 shows consistent results across simulations: shortly after day 61 (the time of the primary  
 390 intervention), we consistently see a decrease in the number of observed cases. The variability in  
 391 the simulations can partly be attributed to the randomness in the stochastic model as well as the  
 392 over-dispersion parameter. Variability can also be explained by the MLE of the basic reproduction  
 393 number being below 2, which together with the stochasticity built into the simulations, can result  
 394 in absence of outbreak, such as in simulation 8, or much smaller outbreaks like in 5, 7, and 9.

395 The MLE of the parameters, which we obtained by picking the maximum of the log  
 396 likelihood over the 100 SMC runs, falls outside the bootstrap 95% CI for  $q$ ,  $\rho_H$ , and  $\psi_H$ . However,  
 397 the distance between the MLE and the boundary of the CIs is small in these three cases, and we  
 398 also run simulations using the bootstrap mean, and all results remained unchanged.

399

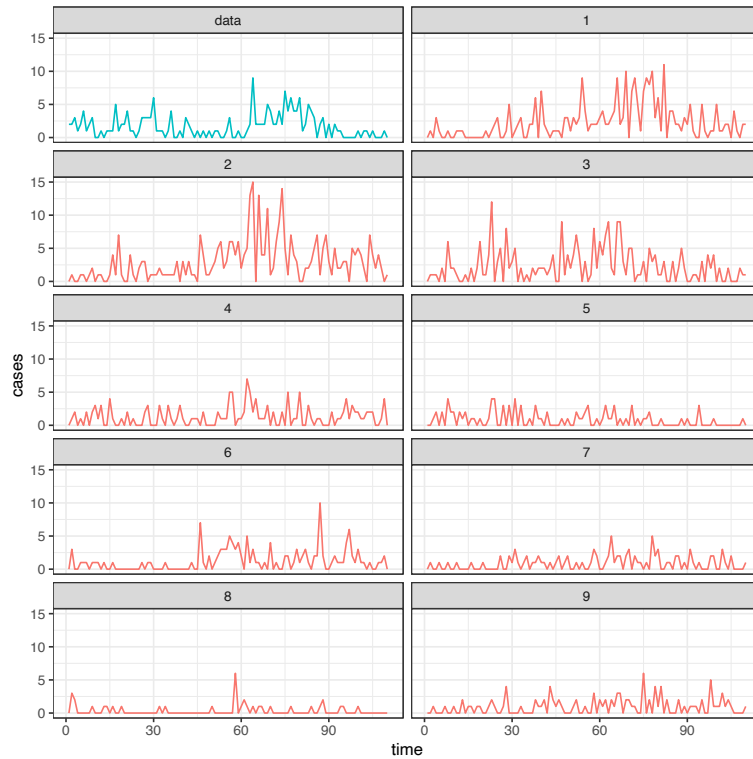
### 400 3.1.2 Maximum Likelihood Estimates for OSU

401 The MLEs of the parameters for the OSU model, as well as derived quantities, are shown in Table  
 402 3. Here we can see an initial reproductive number of almost 6, much higher than Harvard’s.  
 403 However, it eventually becomes lower than 1, which supports our modeling assumptions of an  
 404 awareness campaign from OSU, perhaps helped by news reporting about the outbreak, that lead to  
 405 effective self-isolation of individuals.

<i>Symbol</i>	<i>Description</i>	<i>Point estimate</i>	<i>95% CI</i>	<i>Units</i>	<i>Source</i>
$\beta_O$	Transmission rate constant	1.19	(1.19, 1.2)	day <sup>-1</sup>	MLE
$w$	Decrease in infection due to self-isolation	0.16	(0.157, 1.16)	—	MLE
$\rho_O$	Proportion of infections reported	0.03	(0.029, 0.03)	—	MLE
$\psi_O$	Overdispersion parameter	0.38	(0.376, 0.38)	—	MLE
$R_E(t)$	Effective reproduction number	5.95 initial 0.95 after advisory		—	Calculated as $\frac{\beta(t)}{\gamma(t)}$

406 **Table 3:** List of parameters in the OSU model that obtained by MLE or calculated using  
 407 the estimated parameters.

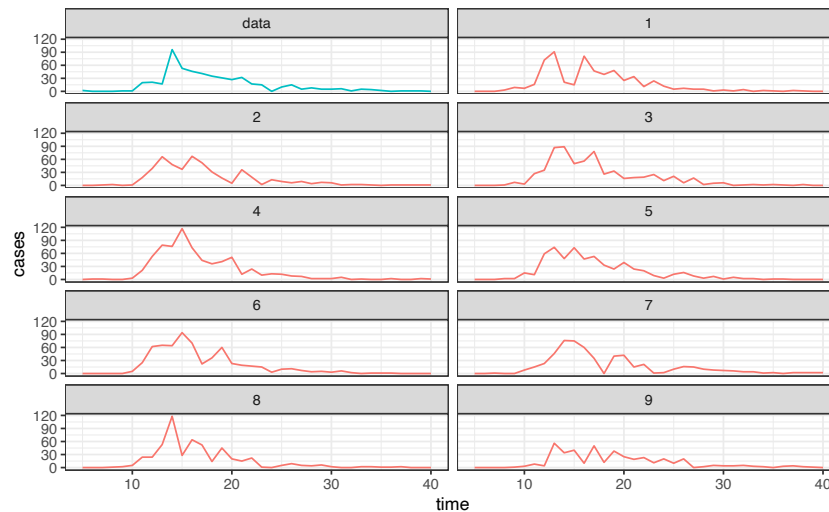
408 As with Harvard, we run stochastic simulations of OSU's outbreak using the parameter values  
409 from Table 3. The simulated outbreaks are shown in Figure 4, and they replicate the real curve  
410 with some random variation due to the stochastic nature of the model.



411

**Figure 3:** Nine simulations of the final Harvard model evaluated at the maximum likelihood estimates. Comparisons to the actual data show that many of the simulations (particularly Simulation 1, 2, 3, and 6) have similar patterns that mirror the shape curve for the observed data.

412



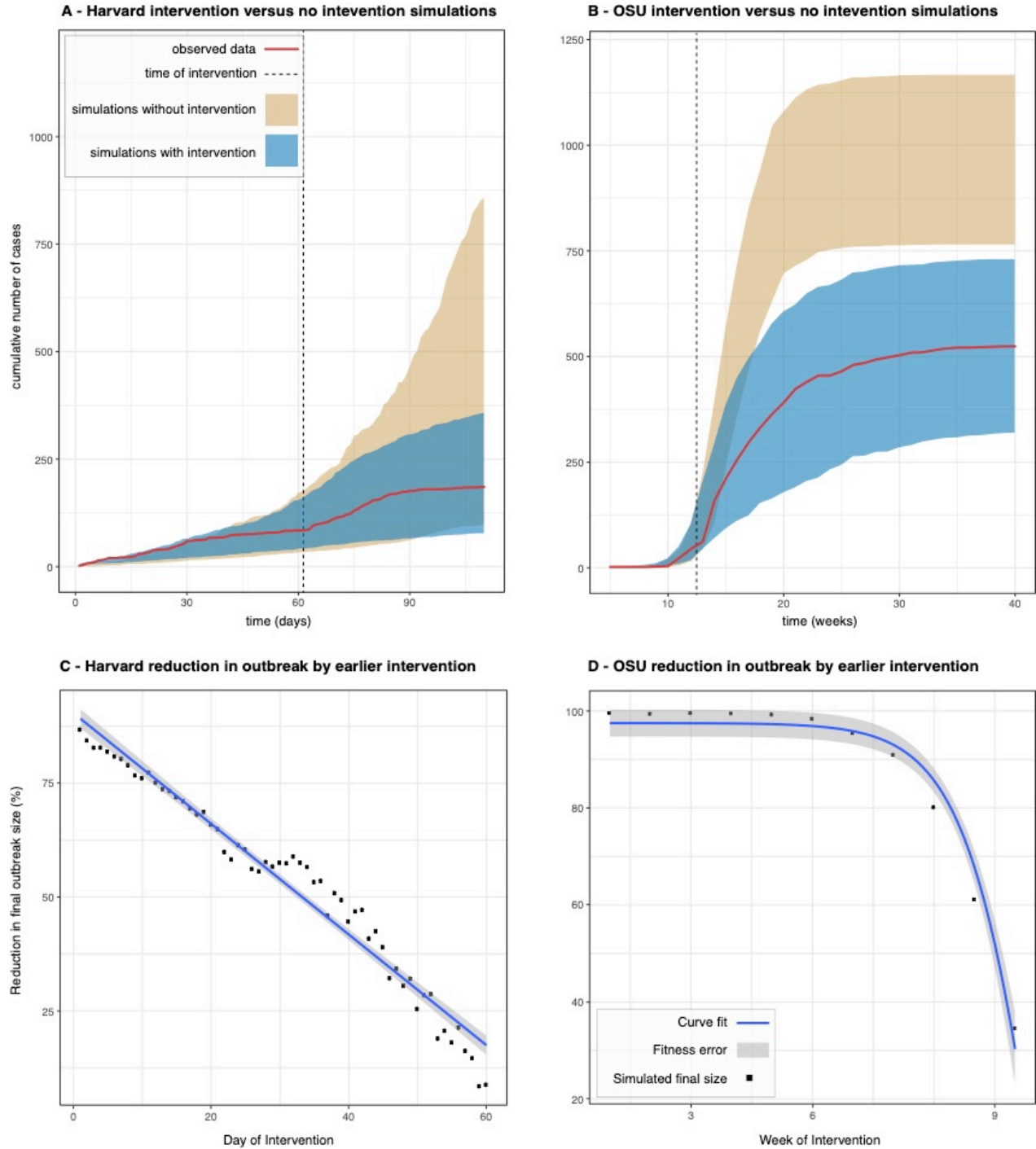
413

**Figure 8:** Nine simulations of the final OSU model evaluated at the maximum likelihood estimates. All of them follow the observed data quite closely.

### 414 **3.2 Earlier intervention decreases outbreak size at Harvard and OSU**

415 The results from the intervention analysis for Harvard and OSU is depicted in Figure 5. By the  
416 final day of the Harvard outbreak (day 130), the simulations without the intervention on day 61  
417 yielded outbreak sizes that were up to four times the size of the actual outbreak (Figure 5A). These  
418 results also indicate that the outbreak would have lasted much longer, if not for these vigilance-  
419 increasing strategies. By varying the day of the intervention from 1 to 61, we also obtained a linear  
420 regression between day of intervention and reduction of the outbreak (Figure 5C). The fitness of  
421 the regression is very high ( $R^2=0.96$ ,  $P<10^{-9}$ ), and quick inspection of the plot reveals that if the  
422 new diagnosis protocol had been implemented within the first 10 days of the outbreak, then no  
423 more than 50 students would have been infected in total at Harvard.

424 For OSU we observe similar trends. Lack of intervention on week 12 could have resulted  
425 in an outbreak twice as large (Figure 5B). The outbreak size as a function of the intervention week  
426 also shows a strong dependency, but in this case non-linear and best fit with a sigmoid function of  
427 the form  $1/(1+e^{\text{week}-12})$ . Using this transformation, the fit is also very high ( $R^2=0.63$ ,  $P<0.005$ ), and  
428 we can conclude that intervening earlier would have had a major effect as well: if the awareness  
429 campaigns prompting students to self-isolate had started around week 5 or 6 (rather than week 12),  
430 then it appears likely that the outbreak could have been completely eradicated.



431

432

433

434

**Figure 5:** Panels A and B show the comparison of the cumulative number of cases over time for the observed Harvard and OSU data and the range of cases (95% percentile of the runs) in simulations with and without interventions, with dotted lines representing the timing of the interventions in each school (panels A and B). In panels C and D, the plots show the percentage we expect outbreak size to decrease by if the date of intervention had been moved up. There is a significant linear relationship between the time and percentage reduction in the case of Harvard, as well as a significant relationship after doing a sigmoid transformation of the time variable in the case of OSU.

## 435 **4. DISCUSSION**

### 436 **4.1 Parameter interpretation**

437 The MLEs give us insight into characteristics of the mumps outbreaks at Harvard University in  
438 2016 and Ohio State University in 2014, as measured by their effective reproduction numbers  $R_E$ ,  
439 intervention parameters  $q$  and  $w$ , rates of removal  $\gamma$ , reporting rates  $\rho$ , and overdispersion  
440 parameters  $\psi$ . At Harvard,  $R_E$  during normal class term was 1.63, which indicates that the  
441 outbreak was growing, even though testing and isolation by HUHS resulted in a baseline removal  
442 time of only  $\frac{1}{0.85} = 1.2$  days. This points to the effectiveness of the quarantine system  
443 implemented by HUHS. However, a small fraction of false negative cases still managed to escape  
444 quarantine and keep the virus under circulation, as indicated by the reproduction number being  
445 higher than 1. The reproduction number goes below 1 during the spring break, which is reasonable  
446 given that most students are away due to the residential nature of the Harvard campus. However,  
447 transmission resumes after the break. It is only after the implementation of the new diagnosis  
448 protocol on day 61, which required isolation if clinical symptoms were present, that had a dramatic  
449 effect on the detection and isolation of positive cases, effectively taking the removal time to less  
450 than 1 day and the reproductive number below 0.6. Thanks to this key intervention, it was possible  
451 to end the outbreak before the beginning of the summer recess.

452 The estimate of  $\rho$  is 0.96, which implies the reporting rate at Harvard was remarkable.  
453 Reasons include the email awareness campaign, a community network – from resident deans to  
454 athletic coaches – reporting students and employees who seemed at-risk, and more aggressive  
455 diagnoses, particularly towards the end of the outbreak. The estimate for  $\psi$  is 0.54, suggesting that  
456 the actual data has more variability than expected under the assumed distribution. If  $\psi$  had been



457 approximately 0, the variance in our measurement model would have simplified to the variance  
458 for a binomial distribution. However, because the 95% confidence interval is (0.5, 0.56) and thus  
459 does not include 0, we justify the modelling decision of representing the number of cases as an  
460 over-dispersed binomial. Demographic and environmental stochasticity (e.g.: a student in the  
461 midst of midterm season may be less likely to report symptoms), as well as the interventions  
462 themselves (e.g.: reporting may increase temporarily after an awareness email) can result in over-  
463 dispersion in the number of reported cases.

464         In the case of OSU, we obtain a much higher reproduction number at the beginning of the  
465 outbreak, near 6, and a very low reporting rate of 3%. Before discussing these results any further,  
466 it is important to keep in mind that we extrapolated OSU cases from state-level reports by the CDC.  
467 Furthermore, we did not have direct access to information about the containment interventions  
468 adopted by the school, as we did for Harvard, so we were only able to make educated guesses  
469 about those possible interventions based on information we found on the web. However, the  
470 internal consistency of the resulting model and the good agreement with the available data, gives  
471 weight to these results. Within our OSU model, we can conclude that self-isolation of students  
472 motivated by the advisories posted by OSU had the intended effect of stopping the outbreak. The  
473 effective reproduction number dips below 1 after March, which is when the awareness campaign  
474 appeared to have started, and also when the outbreak gained local and national prominence due to  
475 news reporting. The low reporting rate is closer to population-wide estimates of this parameter (6),  
476 and is also compatible with a large, non-residential campus where it is harder to reach out to  
477 students as they live scattered around the city. A consequence of this number is that the outbreak  
478 should have been 30 times larger than observed. Since the observed case count is approximately  
479 500, it follows that the total number of cases could have reached 15,000 individuals, which is still

480 possible given that the number of susceptible within the school's student population is over 30,000.  
481 This is still a very significant number, and it is possible that a large majority of these potential  
482 15,000 cases only had mild symptoms. Furthermore, the modeling approximation of closed SEIR  
483 compartments is probably less accurate for OSU given its non-residential nature: students there  
484 have more opportunity to interact with individuals outside of their school, resulting in additional  
485 transmissions that are not captured by our model, and thus affecting the interpretation of  
486 parameters such as the reporting rate.

487

## 488 **4.2 Effect of strict isolation policy vs self-isolation**

489 Arguably the most critical intervention by HUHS was the isolation requirement for confirmed and  
490 probable mumps cases. By comparing the Harvard and OSU outbreaks, we conclude that the  
491 isolation policy led to a smaller average infectious period for Harvard patients. The MLEs for  
492 Harvard and OSU are different for several parameters, most notably basic reproduction number,  
493 reporting rate, and rate of transition from the infectious to removed class. Firstly, OSU's basic  
494 reproduction number is over four times that of Harvard. Harvard's isolation policy best explains  
495 this difference because it physically prevented infectious persons from causing multiple secondary  
496 infections, thus suppressing the growth of the outbreak. Secondly, OSU's reporting rate is  
497 extremely low, at approximately 3% compared to Harvard's 96%. We do not have access to OSU's  
498 diagnostic procedures nor do we know the extent of their email awareness campaign, but we  
499 hypothesize that a lack of one or both of these may explain at least a portion of the dissimilarity in  
500 the two schools' reporting rates. However, the decrease in OSU's transmission rate we observe in  
501 our model post-intervention is still extremely significant with a sixth-fold reduction, and would  
502 have been a major contributor to help containing the outbreak there. This suggests that compliance

503 with easy-to-implement measures such as self-isolation could go a long way towards outbreak  
504 mitigation. Of course, high compliance is contingent on effective educational and awareness  
505 campaigns by the health authorities.

506

### 507 **4.3 Implications of intervention analysis**

508 With the benefit of our intervention analysis, we conclude that aggressive diagnoses decreased the  
509 size of the Harvard outbreak by approximately three-fourths. Furthermore, for every day of  
510 intervention delay, we estimate that the outbreak size would have increased by 1.6 percentage  
511 points, extrapolating the regression line in Figure 5C. Likewise, self-isolation prompted by health  
512 advisories posted by the university reduced the size of the OSU outbreak by half. Given the non-  
513 linear dependency between change in outbreak size and timing of intervention (Figure 5D), the  
514 increase would have been even larger in that outbreak. Interestingly, this dependency also implies  
515 that self-isolation in the first weeks of the outbreak can be enough to completely stop spread.

516 Clearly, a limitation of this analysis is the assumption that everything remains the same  
517 while changing the time of the intervention under consideration. In reality, other factors might  
518 come into play if the outbreak becomes larger or smaller, which in turn could affect the dynamics  
519 of the outbreak as well as the interventions themselves. However, this analysis still provides a  
520 useful hypothetical quantification of the effect of accelerating or delaying interventions designed  
521 to contain the spread of an outbreak and here, as expected, the sooner the interventions are  
522 introduced, the better the outcomes in terms of outbreak size. Of course, existing constraints in the  
523 school's health system could impede fast interventions. In such situations, our method can be

524 useful to perform a cost-benefit analysis of how late an intervention could be made to still have a  
525 significant reduction in the health burden caused by the disease.

526

## 527 **4.4 Conclusions**

528 We constructed and parametrized a POMP model for the transmission of mumps on college  
529 campuses. Unlike other models of infectious disease, which opt for deterministic representations,  
530 our stochastic model is adaptable to small populations and accounts for the noisiness and  
531 incompleteness of case data. Moreover, it incorporates parameters that measures the effect of  
532 interventions implemented after a given point in time. Given the worldwide crisis caused by the  
533 COVID-19 pandemic, such models can be useful to quickly evaluate interventions designed to  
534 contain the spread of SARS-CoV-2 once schools reopen in the U.S. and around the world.

535 We compared an outbreak at Harvard University, with its various intervention strategies,  
536 to another university outbreak of comparable reported cases at OSU. Importantly, while most  
537 literature today focuses on mumps prevention – such as administering third MMR doses to college-  
538 age students – this paper provides quantitative backing for more immediate and less costly  
539 approaches to mitigating the spread of mumps and other infectious diseases, most notably COVID-  
540 19. Even with widespread availability of vaccines, outbreaks of highly transmissible diseases are  
541 still a reality, as mumps exemplifies very clearly. In particular, requiring strict isolation if any  
542 symptoms of the disease are presented would significantly reduce transmission and ultimately the  
543 size of the outbreak. Effective awareness campaigns that lead to self-isolation of infected  
544 individuals with mild symptoms can also have a significant effect in containing the spread of  
545 disease and limiting the risk for vulnerable populations.

## 546 **4.4 Limitations**

547 Some of our conclusions are likely affected by confounding factors that we cannot control for in  
548 this analysis. For example, the outbreak at Harvard started to subside in late April, not long before  
549 students finish the semester and leave campus, which would decrease the number of potential  
550 infections. The most promising method to determine the exact effect of isolation strategies is  
551 through a randomized control trial. Regarding the differences between OSU and Harvard  
552 parameters, we must be cautious in taking the OSU estimates at face value. Given that the OSU  
553 data consists of weekly reports rather than daily reports of cases, we should expect the estimates  
554 for the parameters to be less accurate. Furthermore, the cases are not solely linked to the university.  
555 Numerous cases in the data occurred in the greater Columbus area, suggesting that the parameter  
556 estimates do not only account for the dynamics of mumps on campus. Lastly, major differences in  
557 housing and campus characteristics could have also contributed to differences between the two  
558 schools; for instance, OSU's population size is three times that of Harvard, and OSU has larger  
559 dorms than Harvard's houses. Interventions used at Harvard simply may not have worked as well  
560 at OSU. We were fortunate to have direct access to school administrators who were involved in  
561 the response to the 2016 outbreak to discuss HUHS interventions in detail, but we were not able  
562 to get the same level of detail for OSU's interventions, as discussed in the main text. More broadly,  
563 lack of publicly available datasets, with the exception of CDC reports on OSU's outbreak, is a  
564 serious impediment to perform these analyses. Therefore, it will be essential that universities  
565 across the US and globe actively share data for comparative analysis, to identify the best  
566 intervention strategies to protect college campuses from outbreaks, especially in the post-COVID-  
567 19 world.

568

569 **Competing Interests:** We declare no competing interests.

570 **Source Code:** Available at <https://github.com/broadinstitute/mumps-pomp-models>

571 **Author's Contributions:** MS participated in the design of the study, carried out the data analysis,  
572 developed the epidemiological models, generated the conclusions, and drafted the manuscript; GF  
573 developed the epidemiological models, and generated the conclusions; AC conceived of the study,  
574 participated in the design of the study, coordinated the study, and helped draft the manuscript; SF  
575 and PJB provided data on the HUHS interventions and reviewed the final draft of the manuscript;  
576 PCS overviewed the study and reviewed the final draft of the manuscript. All authors gave final  
577 approval for publication.

578 **Acknowledgements:** The authors would like to thank Jonathan Grad and Joseph Lewnard for  
579 providing feedback on the study design, Hayden Metsky for reviewing the manuscript, members  
580 of MDPH for providing access to the Harvard data, and Bridget Chak and Shirlee Wohl for guiding  
581 in the interpretation of the data.

582 **Ethics:** Usage of Harvard University data for development of the SEIR model was approved by  
583 the Massachusetts Department of Public Health (MDPH) through protocol 906066. Harvard  
584 University Faculty of Arts and Sciences and the Broad Institute ceded review of secondary analysis  
585 to the MDPH IRB through institutional authorization agreements. The MDPH IRB waived  
586 informed consent given this research met the requirements pursuant to 45 CFR 46.116 (d). Data  
587 from Ohio State University was obtained from the CDC's Morbidity and Mortality Weekly Report  
588 2014. The Broad Institute has determined usage of this data constitutes non-human subjects  
589 research.

590 **Funding:** Howard Hughes Medical Institute, US National Institutes of Health.

## REFERENCES

- 591 1. Impact of the COVID-19 pandemic on education Wikipedia, the free encyclopedia2020  
592 [updated 30 May 2020. Available from: [https://en.wikipedia.org/wiki/Impact\\_of\\_the\\_COVID-](https://en.wikipedia.org/wiki/Impact_of_the_COVID-19_pandemic_on_education)  
593 [19\\_pandemic\\_on\\_education](https://en.wikipedia.org/wiki/Impact_of_the_COVID-19_pandemic_on_education).
- 594 2. Panovska-Griffiths J, Kerr CC, Stuart RM, Mistry D, Klein DJ, Viner RM, et al.  
595 Determining the optimal strategy for reopening schools, the impact of test and trace interventions,  
596 and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study.  
597 *Lancet Child Adolesc Health*. 2020;4(11):817-27.
- 598 3. Lordan R, FitzGerald GA, Grosser T. Reopening schools during COVID-19. *Science* (New  
599 York, NY). 2020;369(6508):1146.
- 600 4. Costill D. College campus outbreaks require timely public health response. *Healio*  
601 [Available from: <https://goo.gl/aW7GaT>.
- 602 5. Hurtado C. The Rise of Mumps in the United States Outbreak Observatory [Available from:  
603 [https://www.outbreakobservatory.org/outbreakthursday-1/12/7/2017/the-rise-of-mumps-in-the-](https://www.outbreakobservatory.org/outbreakthursday-1/12/7/2017/the-rise-of-mumps-in-the-united-states)  
604 [united-states](https://www.outbreakobservatory.org/outbreakthursday-1/12/7/2017/the-rise-of-mumps-in-the-united-states).
- 605 6. Lewnard JA, Grad YH. Vaccine waning and mumps re-emergence in the United States.  
606 *Science Translational Medicine*. 2018;10(433).
- 607 7. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and  
608 Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med*. 2020.
- 609 8. The COVID vaccine challenges that lie ahead. *Nature*. 2020;587(7835):522.
- 610 9. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung S-m, Hayashi K, et al. Estimation of  
611 the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of*  
612 *Infectious Diseases*. 2020;94:154-5.
- 613 10. Davies N, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, Pearson CAB, et al.  
614 Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in  
615 England. *CMMID Repository*. 2020.
- 616 11. Shah M, Quinlisk P, Weigel A, Riley J, James L, Patterson J, et al. Mumps Outbreak in a  
617 Highly Vaccinated University-Affiliated Setting Before and After a Measles-Mumps-Rubella  
618 Vaccination Campaign—Iowa, July 2015–May 2016. *Clinical Infectious Diseases*. 2018;66(1):81-  
619 8.
- 620 12. Golwalkar M, Pope B, Stauffer J, Snively A, Clemmons N. Mumps Outbreaks at Four  
621 Universities - Indiana, 2016. *MMWR Morb Mortal Wkly Rep*. 2018;67(29):793-7.
- 622 13. Bixler J, Botelho G. Over 360 cases of mumps in central Ohio, most of them tied to OSU.  
623 *CNN* [Available from: <https://www.cnn.com/2014/05/16/health/ohio-mumps/index.html>.
- 624 14. Mumps Cases and Outbreaks Centers for Disease Control and Prevention [Available from:  
625 <https://www.cdc.gov/mumps/outbreaks.html>.
- 626 15. Wohl S, Metsky HC, Schaffner SF, Piantadosi A, Burns M, Lewnard JA, et al. Combining  
627 genomics and epidemiology to track mumps virus transmission in the United States. *PLOS*  
628 *Biology*. 2020;18(2):e3000611.
- 629 16. Isolation Expenses (2016-2017). *Harvard University Health Services*; 2017.
- 630 17. Stocks T, Britton T, Höhle M. Model selection and parameter estimation for dynamic  
631 epidemic models via iterated filtering: application to rotavirus in Germany. *Biostatistics*. 2018.
- 632 18. Barreira P, Fitzgerald S. Personal communication. 2018.

- 633 19. Correspondence regarding Mumps 2016, Spring to Summer. Harvard University Health  
634 Services.
- 635 20. Mumps List for CDC. Harvard University Health Services.
- 636 21. Massachusetts Mumps Case Data between 2015 and 2017. Massachusetts Department of  
637 Public Health.
- 638 22. Bitsko RH, Cortese MM, Dayan GH, Rota PA, Lowe L, Iversen SC, et al. Detection of  
639 RNA of mumps virus during an outbreak in a population with a high level of measles, mumps, and  
640 rubella vaccine coverage. *J Clin Microbiol.* 2008;46(3):1101-3.
- 641 23. Morbidity and Mortality Weekly Report. 2014.
- 642 24. Mumps Outbreak FAQs: The Ohio State University; [Available from:  
643 <https://www.osu.edu/assets/mumps/faqs.pdf>.
- 644 25. The Ohio State University Mumps Outbreak FAQs Columbus Public Health; 2014  
645 [Available from: [http://www.ccchd.com/documents/contentdocuments/doc\\_23\\_5\\_516.pdf](http://www.ccchd.com/documents/contentdocuments/doc_23_5_516.pdf).
- 646 26. Zeltner B. Mumps outbreak at Ohio State: Five things you need to know  
647 cleveland.com2014 [updated January 12 2019. Available from:  
648 [https://www.cleveland.com/healthfit/2014/03/mumps\\_outbreak\\_at\\_ohio\\_state\\_f.html](https://www.cleveland.com/healthfit/2014/03/mumps_outbreak_at_ohio_state_f.html).
- 649 27. Martcheva M. *An Introduction to Mathematical Epidemiology*: Springer; 2015.
- 650 28. King AA, Nguyen D, Ionides EL. Statistical Inference for Partially Observed Markov  
651 Processes via the R Package pomp. *Journal of Statistical Software.* 2016;69(12):43.
- 652 29. Brauer F, Van den Driessche P, Wu J. *Mathematical epidemiology*: Springer; 2008.
- 653 30. He D, Ionides EL, King AA. Plug-and-play inference for disease dynamics: measles in  
654 large and small populations as a case study. *Journal of The Royal Society Interface.*  
655 2010;7(43):271-83.
- 656 31. Academic Calendar: 2015-16: Harvard Law School; 2015 [Available from:  
657 <https://hls.harvard.edu/dept/academics/academic-calendar/academic-calendar-2015-16/>.
- 658 32. The Ohio State University - Statistical Summary: The Ohio State University; 2013  
659 [Available from: <https://www.osu.edu/osutoday/stuinfo13.php>.
- 660 33. King AA, Ionides EL, Bretó MC, Ellner SP, Ferrari MJ, Kendall BE, et al. pomp: Statistical  
661 Inference for Partially Observed Markov Processes. 2020.
- 662 34. R Core Team. *R: A Language and Environment for Statistical Computing.* 2019.
- 663 35. Davison AC, Hinkley DV. *Bootstrap Methods and Their Applications*: Cambridge  
664 University Press; 1997.
- 665 36. Canty A, Ripley BD. boot: Bootstrap R (S-Plus) Functions. 2017.