Received: 19 June 2020 | Revised: 29 September 2021 | Accepted: 5 October 2021

DOI: 10.1111/biom.13579

Biometrics WILEY

BIOMETRIC PRACTICE

An individual level infectious disease model in the presence of uncertainty from multiple, imperfect diagnostic tests

Caitlin Ward | **Grant D. Brown** | **Jacob J. Oleson**

Department of Biostatistics, University of Iowa, Iowa City, Iowa, USA

Correspondence
Caitlin Ward, Department of Biostatistics, University of Iowa, Iowa City, IA 52242, USA.
Email: caitlin-e-ward@uiowa.edu

Funding information
Fogarty International Center, Grant/Award Number: R01TW010500

Abstract

Bayesian compartmental infectious disease models yield important inference on disease transmission by appropriately accounting for the dynamics and uncertainty of infection processes. In addition to estimating transition probabilities and reproductive numbers, these statistical models allow researchers to assess the probability of disease risk and quantify the effectiveness of interventions. These infectious disease models rely on data collected from all individuals classified as positive based on various diagnostic tests. In infectious disease testing, however, such procedures produce both false-positives and false-negatives at varying rates depending on the sensitivity and specificity of the diagnostic tests being used. We propose a novel Bayesian spatio-temporal infectious disease modeling framework that accounts for the additional uncertainty in the diagnostic testing and classification process that provides estimates of the important transmission dynamics of interest to researchers. The method is applied to data on the 2006 mumps epidemic in Iowa, in which over 6,000 suspected mumps cases were tested using a buccal or oral swab specimen, a urine specimen, and/or a blood specimen. Although all procedures are believed to have high specificities, the sensitivities can be low and vary depending on the timing of the test as well as the vaccination status of the individual being tested.

KEYWORDS

Bayesian, compartmental model, diagnostic uncertainty, infectious disease modeling, reversible-jump MCMC

1 | INTRODUCTION

Mathematical modeling of infectious diseases allows researchers to study the complex dynamics of disease transmission and the effectiveness of interventions, which are of practical importance to public health organizations attempting to stop the spread of such diseases. These models offer essential tools to understand epidemiological patterns, and accurate modeling gives researchers increased knowledge of disease spread to help inform policy decisions on disease control measures (Heesterbeek *et al.*, 2015). However, these models often rely on imperfect diagnostic tests to identify individuals with the disease despite the presence of test uncertainty, leading to incorrect results.

Compartmental models capture the dynamics of infectious diseases by parameterizing the probabilities of transitions between important disease states. The most common such model is the SIR model, which has three compartments: **S**usceptible, **I**nfectious, and **R**emoved. These models can be implemented deterministically using ordinary differential equations or stochastically, with the

 © 2021 The International Biometric Society

primary method for inference using Bayesian methodology. There are two primary categories of stochastic compartmental models: population-averaged and individual-level, with population-averaged models assuming all individuals within the population or sub-population have the same transition probabilities, and individual-level models allowing these probabilities to be influenced by individual characteristics.

Typical infectious disease data record either individual diagnosis dates, or counts of positive diagnoses over time. However, in many cases, these diagnoses are subject to uncertainty from one or multiple imperfect diagnostic tests, (e.g., Evans *et al.*, 2010; Akhtar and Carpenter, 2013; Porter and Oleson, 2013). Currently, infectious disease models generally assume that subjects with positive diagnostic test results are infectious and those with negative test results are not. In other words, these models assume false positive and negative rates of 0. This may be reasonable for highly sensitive and specific tests, but could have severe consequences otherwise. This can be especially challenging for researchers attempting to model newly emerging infectious agents for which lab confirmation may take weeks, but convenient diagnostics with short processing times are not as accurate (Chan *et al.*, 2017).

There is an established literature on diagnostic test inaccuracy in the absence of a gold standard. Latent class models (LCMs) were initially proposed by Hui and Walter (1980) and have been applied in various settings. Collins and Huynh (2014) provide a thorough review of LCMs used in medical applications. LCMs reformulate the probability of obtaining a set of test results in terms of the sensitivity and specificity of the diagnostics, as well as the population prevalence of the disease, allowing for inference on the diagnostic characteristics. LCMs can be fit with maximum likelihood approaches, but more recent extensions have focused on Bayesian models (Gastwirth *et al.*, 1991). The Bayesian framework offers the advantage of including informative prior information about known diagnostic characteristics.

In this paper, we present a novel individual-level model (ILM) formulation that incorporates diagnostic testing uncertainty. The diagnostic ILM blends ideas from traditional ILMs and LCMs to model disease transmission and latent disease status simultaneously. We also present a new method for estimating the infectious period in the ILM framework. We show via simulation that accounting for diagnostic characteristics yields better estimates of the transmission probability. To account for the additional model complexity from estimating disease status, we describe a method for accurate and efficient posterior sampling of infectiousness status. Finally, the model is applied to data from the Iowa 2006 mumps epidemic.

TABLE 1 Test results from the Iowa mumps epidemic

		Swab		
Urine	**Serum**	**Positive**	**Negative**	**No test**
Positive	Positive	1	7	1
	Negative	3	9	1
	No test	1	0	1
Negative	Positive	29	609	17
	Negative	61	1,983	66
	No test	9	204	31
No test	Positive	35	539	233
	Negative	56	1,076	858
	No test	19	233	0

2 | METHODS

2.1 | Motivating data

The data motivating the development of the diagnostic ILM come from the 2006 Iowa mumps epidemic (Marin *et al.*, 2008) and consists of 6,082 individuals that were tested for mumps between January 3, 2006 and June 23, 2006. All diagnostic testing for this data set was done at the State Hygienic Laboratory at the University of Iowa. Two diagnostics are recommended by the CDC to detect mumps in patients with compatible symptoms: a buccal or oral swab specimen for detection of viral RNA by RT-qPCR, and a serum specimen for IgM detection (CDC, 2019). In some cases, urine specimens are also tested by RT-qPCR, particularly if patients are presenting with mumps complications. We shall hereafter refer to the three diagnostics simply as the swab, serum, and urine tests. Of the 6,082 subjects tested, 2,702 were tested with all three diagnostics, 1,706 received only the serum and swab tests, and 1,091 received only the serum test. Table 1 summarizes the test results for all subjects included in this analysis. Previous infectious disease modeling of this epidemic used only the 214 positive swab tests (Porter and Oleson, 2013).

Mumps has an incubation period of 16–18 days, with extremes of 12–25 days (Hamborsky *et al.*, 2015). Because of this, mumps epidemics are well approximated using SEIR compartmental models that add a latent **E**xposed compartment to the traditional SIR model (e.g., Chen *et al.*, 2007; Porter and Oleson, 2013; Trotz-Williams *et al.*, 2017). The infectious period is less well-known, but is often between 5 and 9 days (Polgreenet *et al.*, 2008). This epidemic occurred in a primarily vaccinated population, with estimates indicating 6% of patients were unvaccinated, 12% had one dose of the measles, mumps, and rubella (MMR) vaccine, 51% had received 2 doses, and 31% did not have vaccination records (CDC, 2006). Evidence suggests that immunized patients have a different serological reaction to the

infection, causing the sensitivities of the tests to be as low as 30–50% (Sanz *et al.*, 2006; Peltola *et al.*, 2007; Brockhoff *et al.*, 2010; Rota *et al.*, 2013). With such low sensitivities, we can expect many false negative results. Incorporating diagnostic testing uncertainty in the transmission model is therefore important.

In addition to the results of three diagnostics tests, the mumps epidemic data also included symptom onset date, date of specimen collection, city of residence, age, and sex for each tested individual. Symptom onset date was missing for 31.8% of individuals, but specimen collection date was always recorded. The data do not differentiate between tested individuals without symptoms and symptomatic individuals with missing symptom onsets. Due to this, we assume symptom onset dates are missing at random and used a linear regression to impute these dates from the collection date prior to the main analysis.

2.2 | Diagnostic ILM model

The diagnostic ILM framework can be incorporated into various compartmental structures (e.g., SIR, SEIR, and SIS), but we present it as SEIR model for generality. Let $i = 1, \dots, N$ index individuals and $t = 0, \dots, \tau$ denote discrete calendar time since the beginning of the epidemic to recovery (or removal) of the last infectious individual. Let $\mathbf{Z^S}, \mathbf{Z^E}, \mathbf{Z^I}$, and $\mathbf{Z^R}$ each be $N \times \tau$ matrices that describe the state of each individual at each time-point. The itth element of these matrices, Z_{it}^c for $c \in \{S, E, I, R\}$ is either a 0 or a 1, which indicates the infection state of the ith individual at time t. As an individual cannot be in more than one compartment at any one time, $\mathbf{Z^S} + \mathbf{Z^E} + \mathbf{Z^I} + \mathbf{Z^R} = \boldsymbol{\mathcal{J}}_{N\times\tau}$, where $\boldsymbol{\mathcal{J}}_{N\times\tau}$ is a matrix of ones.

Diagnostic testing results are incorporated in the data model. We assume dichotomous results on each test and specify Y_{ik} to contain individual i's test results on all tests, k, where $k = 1, \dots, K_i$. Let η_i represent the latent disease status with $\eta_i = 1$ indicating true disease status and $\eta_i = 0$ indicating otherwise. In terms of the matrices defined previously, $\eta_i = I(\sum_{t=0}^{\tau} Z_{it}^I \geq 1)$. Assuming that the test results are independent given the latent disease status, the likelihood can be written as $P(\mathbf{Y}_i = \mathbf{y}_i|\eta_i) = \prod_{k=1}^{K_i} P(Y_{ik} = y_{ik}|\eta_i)$, where $P(Y_{ik} = y_{ik}|\eta_i)$ is a function of the sensitivity or specificity of test k, depending on the underlying disease status $\eta_i \in \{0, 1\}$.

Process models describe the transitions between compartments. Let Z_{it}^{*E}, Z_{it}^{*I}, and Z_{it}^{*R} denote individual indicators for transitioning to the superscripted compartment in the interval $(t, t+1]$. The timing of these transitions depends on observed data. Most commonly, symptom onset date will be used to inform the time at which $Z_{it}^{*I} = 1$ for individuals with $\eta_i = 1$, and the timing of the other transitions depends on the estimated length of time spent in each compartment. These components are assigned a chain Bernoulli structure: $Z_{it}^{*E} \sim Bern(\pi_{it}^{(SE)})$, $Z_{it}^{*I} \sim Bern(\pi_{it}^{(EI)})$, and $Z_{it}^{*R} \sim Bern(\pi_{it}^{(IR)})$.

We use the transition probability defined by Deardon *et al.* (2010) to describe the infection process, with a modification allowing transmissibility to change over epidemic time. The probability is written as

$$\pi_{it}^{(SE)} = 1 - \exp\left[\left\{-\Omega_S(i)\sum_j I(Z_{jt}^I = 1)\Omega_T(j,t)\kappa(i,j)\right\} - \epsilon\right], \tag{1}$$

where $\Omega_S(i)$ describes risk factors associated with susceptibility of individual i, $\Omega_T(j,t)$ describes risk factors associated with transmissibility of infectious individual j at time t, $\kappa(i,j)$ is the infection kernel (usually based on separation distance, d_{ij}), and $\epsilon > 0$ is the sparks term that accounts for infections not explained by the data or the modeling framework. In the diagnostic ILM, the sparks term incorporates transmission from untested infectious individuals and tested infectious individuals that may be misclassified in the model.

We assume individuals are not retested at a later time and that tested individuals who are not classified as infectious do not later contract the disease. As symptom onset date is used as a proxy for the start of infectiousness, individuals with $\eta_i = 1$ transition from **E** to **I** at this time (and from **S** to **E** prior to symptom onset) while individuals with $\eta_i = 0$ transition directly from **S** to **R**. This assumption could be relaxed in practice, however, the motivating data did not include individuals that were tested at multiple times. Additionally, it is plausible that those individuals that were tested but do not have the disease are sick with something else and will modify their behavior to avoid future transmission. For epidemics where individuals are repeatedly tested, this assumption could be relaxed by having tested individuals that are not classified as infectious remaining in **S** with the possibility of transitioning at multiple time points, but this would require estimation of the latent disease status for an individual for each time they were tested.

The transitions from **E** to **I** and **I** to **R** can be fixed or modeled in terms of the transition probabilities. We propose a fixed latent period and an ILM extension of the path-specific approach of Porter and Oleson (2013) to model the infectious period as simultaneous estimation of both periods is challenging with limited data and the latent period for mumps is well known. Using a fixed latent period implies that the timing of the **S** to **E** transition for infectious

individuals is also fully characterized by the observed data. The transition probability is defined as $\pi_w^{(IR)} = P(W \leq w + 1 | W > w)$, where w denotes the discrete time spent in the compartment and W is a random variable defining the infectious period. Common choices of distribution for W are exponential, gamma, or Weibull. The amount of time an individual has spent infectious, $\tau_I(i) = \sum_{t=0}^{\tau} Z_{it}^I$, changes the probability, not epidemic time, so we write the individual contribution to the log-likelihood for the transition from **I** to **R** as $\eta_i[\{\sum_{w=1}^{\tau_I(i)-1} \log(1-\pi_w^{(IR)})\} + \log(\pi_{\tau_I(i)}^{(IR)})]$. $\tau_I(i)$ is unobserved and must be estimated as part of the modeling process. Theoretically, $\tau_I(i)$ can range between 1 and ∞, but it is better to choose an upper bound reflective of knowledge on the disease process, after which $\pi_w^{(IR)} \equiv 1$.

Combining the data and process models, the full log-likelihood can be written as

$$\begin{aligned}
&\log L(\mathbf{Y}, \mathbf{Z}^{*E}, \mathbf{Z}^{*I}, \mathbf{Z}^{*R} | \theta) \\
&= \sum_{i=1}^{N} \Bigg[I(y_i^T = 1) \sum_{k=1}^{K} \log P(Y_{ik} = y_{ik} | \eta_i) \\
&\quad + \sum_{t=0}^{\tau} I(Z_{it}^S = 1) \Big\{ Z_{it}^{*E} \log(\pi_{it}^{(SE)}) \\
&\quad + (1 - Z_{it}^{*E}) \log\left(1 - \pi_{it}^{(SE)}\right) \Big\} \\
&\quad + \eta_i \left[\left\{ \sum_{w=1}^{\tau_I(i)-1} \log\left(1 - \pi_w^{(IR)}\right) \right\} + \log\left(\pi_{\tau_I(i)}^{(IR)}\right) \right] \Bigg],
\end{aligned} \tag{2}$$

where y_i^T is an indicator for individual i getting tested during the epidemic.

3 | COMPUTING

3.1 | Infectious classifications

When fit to large populations, ILMs are computationally challenging and generally rely on MCMC sampling techniques (Deardon *et al.*, 2010). An important consideration in constructing a sampling scheme for the diagnostic ILM is the proposal for η_i. Disease status is latent and must be proposed within the Markov chain and accepted or rejected with probability preserving the stationary distribution. This is not straightforward, as the inclusion or removal of an individual in the infectious state also involves changing the dimension of the collection of infectious periods, that is, if $\eta_i = 0$, then $\tau_I(i)$ does not exist. To account for the changing dimension, we use a reversible-jump MCMC (RJ-MCMC, Green (1995)) step in the proposal of infectious classifications. Technical details about the RJ-MCMC specifications for the diagnostic ILM can be found in Web Appendix A.

3.2 | Infectious period

Another consideration is the proposal for $\tau_I(i)$, as proposing times for all infectious individuals at once leads to poor acceptance. For this reason, we recommend proposing $\tau_I(i)$ for a small proportion of individuals at each MCMC iteration, and found updating around 5% of the individuals tested led to a satisfactory acceptance rate. We investigated proposals with independent draws from the prior distributions for the parameters defining W and a Gaussian random walk, but found that a proposal with a 50% chance of increasing or decreasing the current length by one worked well. One advantage of this scheme is that the proposal can account for the predefined range for $\tau_I(i)$ by mandating that proposals from the endpoints stay in the specified range. Although this approach improves convergence, slow mixing and high autocorrelation remain problematic in the parameters defining W.

4 | SIMULATION STUDY

4.1 | Simulation set-up

To determine the effectiveness of incorporating diagnostic testing uncertainty, we implement a simulation study. Although the simulation is designed to mimic the characteristics of mumps and its associated diagnostics, the assumptions can also be extrapolated to other infectious diseases. The sample size used was $N = 500$ and each individual was given a unique spatial location in one of four Iowa counties with equal probability—Black Hawk, Johnson, Dubuque, or Polk. To incorporate false positive and true negative test results, susceptible individuals were sampled for testing using a logistic model that allows the testing probability to vary with disease prevalence in each of the four counties over the course of the epidemic. If tested, susceptibles then move to the removed compartment as we do not allow for repeat testing.

Although the observed data included three diagnostics, only two were considered in the simulation, with diagnostic characteristics chosen to reflect the serum and swab tests used to classify mumps. Sensitivities of 50% and 59% were chosen for the serum and swab tests, respectively, to reflect sensitivities when performed on samples from vaccinated individuals (Rota *et al.*, 2013). These low sensitivities result in many false negative tests. Many

studies have indicated high specificities (> 99%) for both diagnostics (Krause *et al.*, 2006; Boddicker *et al.*, 2007; Hatchette *et al.*, 2009), so specificities for both tests were set at 99%. For infectious individuals, testing occurred at the time of transitioning from **E** to **I**, and for tested susceptible individuals, it occurred on the day the individual was sampled for testing. Test date is also used as the time of transitioning from **E** to **I** when tested susceptible individuals are misclassified as infectious during model fitting. Untested individuals are assumed to remain susceptible throughout the epidemic.

Transmission was modeled using a simplified form of Equation (1) with $\Omega_S(i)\Omega_T(j,t) = \psi e^{-\omega_T(t-t^*)_+}$ and $\kappa(i,j) = d_{ij}^{-\rho}$, where $(t-t^*)_+ = (t-t^*)I(t \geq t^*)$, resulting in a final form of $\pi_{it}^{(SE)} = 1 - \exp\{-\psi e^{-\omega_T(t-t^*)_+} \times \sum_j I(Z_{jt}^I = 1) d_{ij}^{-\rho}\}$. In this formulation, ψ is an infectivity parameter, and we model constant transmissibility until the time of an exponential decay intervention, t^*. The distance, d_{ij}, between susceptible individual i and infectious individual j is scaled by $\rho > 0$ such that individuals closer together are more likely to transmit disease. Although not an issue in our simulations, in the event that $d_{ij} = 0$, the kernel would become infinite and the transmission probability would be 1. This can be avoided by setting distances of 0 to some pre-specified "small" value, d_0 or incorporating an additional parameter for subjects sharing the same spatial location. The true parameter values in the simulation were $\psi = 0.1$, $\omega_T = 0.07$, and $\rho = 1.5$. These parameters were chosen to simulate epidemics that were not prone to immediately dying out, were similar in length to the observed mumps epidemic ($\sim$ 200 days), and were likely to spread between the counties. The intervention timing was chosen to represent a CDC dispatch in the observed mumps epidemic (Porter and Oleson, 2013). Using the timing of the first positive test as day 1, this intervention occurs at $t^* = 92$.

The latent period was fixed at 17 days, and the infectious period was modeled with a truncated Weibull random variable as described in Section 2.2, with shape parameter $\alpha_I = 7$, scale parameter $\beta_I = 8$, and a maximum number of infectious days of 10. The Weibull distribution was chosen as it was found to provide the best model fit for the infectious duration in the previous model of the observed epidemic (Porter and Oleson, 2013). Susceptible and infectious individuals time of testing and test results were recorded and used to fit the models, with test time informing the transition times. Fifty simulated epidemics were created, each using the same locations and true parameters. Additional details about the simulation set-up can be found in Web Appendix B.

On each simulated epidemic, five different models were fit exploring possible ways to use the testing data. The first model (Reference) assumes we had the true knowledge of infectiousness for all individuals. In practice, when there is diagnostic uncertainty these data are not available, but it provides a useful baseline. The next three heuristic models use the union of the diagnostics (Serum or Swab +), all individuals positive on the serum test (Serum +), and all individuals positive on the swab test (Swab +) to classify individuals as susceptible or infectious prior to model fitting. These three models and the reference model were all fit using the standard ILM framework, which does not incorporate test uncertainty. Therefore, the only difference between analyses is the data used in the model fitting procedure. The final model fitted is the diagnostic ILM model proposed here (Diagnostic ILM), which uses diagnostic test results (positive or negative) from all tested individuals (susceptible or infectious) and models the uncertainty about the infectiousness status.

All models were fit with the transmission and removal probabilities parameterized according to the way they were simulated, except the sparks term (ϵ as defined in Equation 1) was included in the specification of the transmission probability. The sparks term was used only in model fitting to assess the impact of excluding false negative tests in the serum or swab +, serum +, and swab + models, as well as the misclassification of truly infectious individuals in the diagnostic ILM. By simulating data without the sparks term, the estimates for ϵ are characterized only by the choice of model and not by a random infection process. The primary parameters of interest estimated in all models were $\{\psi, \omega_T, \rho, \epsilon, \alpha_I, \beta_I\}$. The diagnostic ILM also estimates the sensitivity and specificity of both diagnostics.

In the Bayesian framework priors must be specified for all parameters. Parameters ψ, ω_T, and ρ were given vague gamma distributions with shape 0.1 and rate 0.1. The sparks parameter ϵ was given a vague half-normal prior with variance 1,000. Since there is prior knowledge on the infectious duration of mumps, α_I and β_I were given loosely informative priors of Gamma(27,3) and Gamma(45,5), respectively. For the diagnostic ILM, we chose informative priors to correspond to what would be done in practice. Priors for the serum test on sensitivity and specificity were Beta(50, 50) and Beta(99, 1), respectively. Priors for the swab test on sensitivity and specificity were Beta(60, 40) and Beta(99, 1), respectively.

4.2 | Simulation results

Three chains were run for each of the five models fit to each of the 50 simulated epidemics. Convergence was assessed with the Gelman–Rubin diagnostic (Gelman *et al.*, 1992). All chains ran for a burn-in of at least 50,000 iterations and

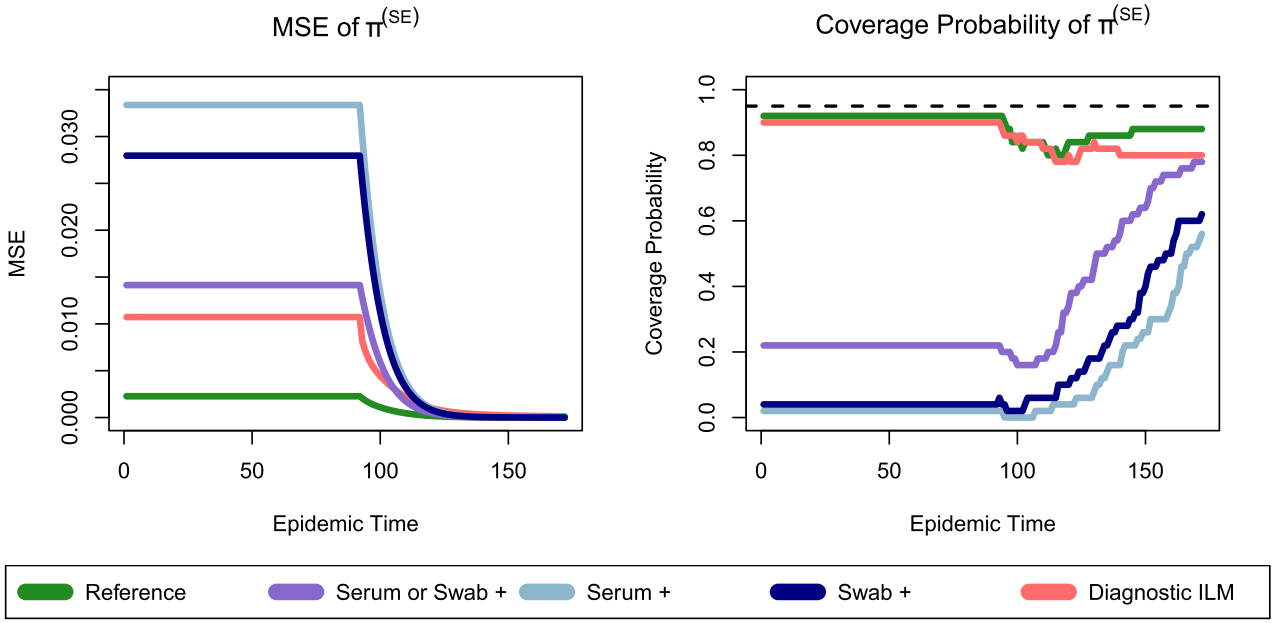

FIGURE 1 MSE and coverage probability of 95% credible intervals for the transmission probability for all models over epidemic time. MSE is calculated as the average of the squared distances between the posterior mean and the true value over 50 simulated epidemics. This figure appears in color in the electronic version of this article, and any mention of color refers to that version

were subsequently run until the Gelman–Rubin diagnostic was less than 1.1 with thinning applied to reduce autocorrelation.

The five models are compared in terms of mean square error (MSE) and the coverage probability of 95% credible intervals constructed for the transmission probability $\pi_{it}^{(SE)}$ and the removal probabilities $\pi_{w}^{(IR)}$. MSE measures the average squared deviations between posterior mean estimates and the true probability. Coverage probability was calculated as the proportion of the 50 simulations where the 95% credible interval contained the true value. The transmission probability was estimated assuming a 0.5 km distance between a susceptible individual and one infectious individual, from day 0 until the maximum epidemic length of any of the simulations, which was 172 days.

Figure 1 displays the MSE and coverage probability of the transmission probability for each model over epidemic time. Performance can be compared to the reference model, which has a small amount of bias due to prior shrinkage. These results highlight several interesting implications. The diagnostic ILM has the lowest MSE prior to the intervention, with the serum or swab + model having slightly more bias until after the intervention occurs. Additionally, the diagnostic ILM vastly outperforms the serum or swab + model in terms of coverage probability, although the serum or swab + model's coverage improves after the intervention. The serum + and swab + models perform the worst in terms of both MSE and coverage probability. The swab diagnostic has slightly higher sensitivity, which improves estimation compared to the serum + model.

Table 2 gives the MSE and coverage probabilities for each of the primary parameters of interest. This shows the relatively poor estimation of the intervention parameter by the diagnostic ILM, but improved estimation and coverage of other transmission parameters. More discussion of the estimation of these parameters is provided in Web Appendix B. Model performance is comparable in estimating the Weibull parameters associated with the removal probability. All models show some bias in these parameters, although with high coverage due to large variance in the estimates. Figure 2 shows the estimates of the transition probabilities over the course of an individual's infectious period using the posterior means from a representative simulation compared to the truth, which illustrates the effects of the bias in the estimation. Credible intervals are provided in Web Table 1. All models tended to favor longer infectious durations, shown by the underestimation of the transition probabilities, especially for transitions after a shorter duration.

This simulation study yields several important findings. When the sensitivities of both diagnostics are low, choosing only one diagnostic to define the infectious class results in substantial bias and low coverage of the transmission probability. Comparatively, using the union of positive tests to define the infectious class results in bias only slightly worse than the diagnostic ILM, but coverage remains very low. In this setting, where diagnostics are likely to produce false negatives, incorporating diagnostic

TABLE 2 MSE (Coverage Probability) for transmission and removal probability parameters. MSE is calculated as the average of the squared distances between the posterior mean and the true value over 50 simulated epidemics. Transmission is a function of ψ (baseline infectivity), ω_T (intervention), and ρ (distance power) and α_I and β_I are the Weibull shape and scale parameters defining the removal probability

	ψ	ω_T	ρ	α_I	β_I
Reference	0.0003 (0.94)	0.0002 (0.88)	0.0063 (0.94)	4.6116 (0.98)	0.1191 (1.00)
Serum or Swab +	0.0020 (0.22)	0.0004 (0.90)	0.0464 (0.30)	4.1731 (0.98)	0.1315 (0.96)
Serum +	0.0050 (0.02)	0.0007 (0.90)	0.1708 (0.14)	4.0081 (1.00)	0.2154 (0.94)
Swab +	0.0041 (0.04)	0.0007 (0.90)	0.1190 (0.14)	4.0561 (1.00)	0.2075 (0.96)
Diagnostic ILM	0.0018 (0.90)	0.1115 (0.82)	0.0356 (0.88)	4.7101 (1.00)	0.2679 (0.96)

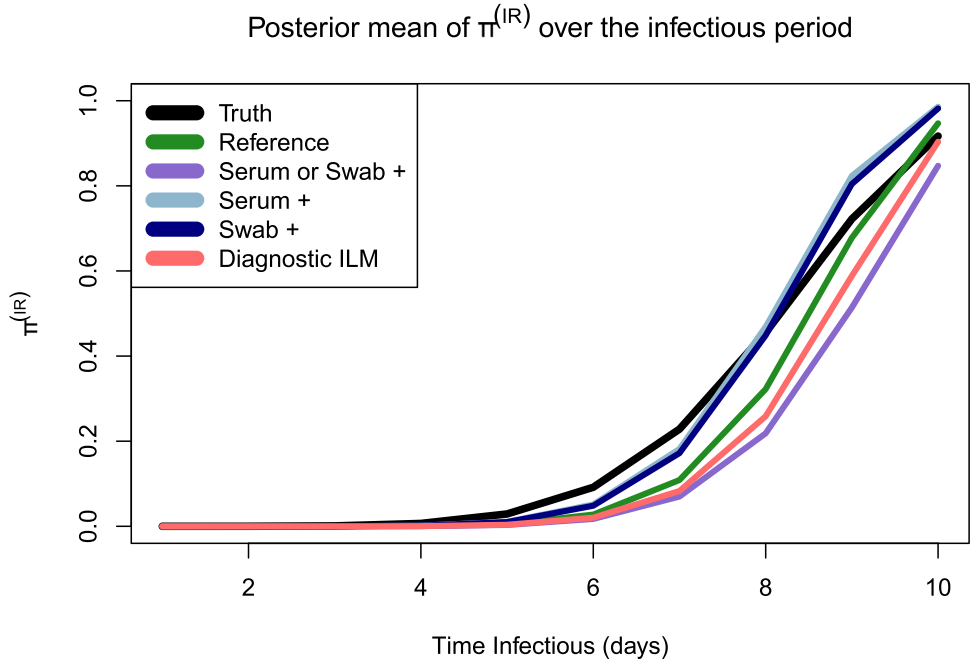

FIGURE 2 Posterior mean estimates of removal transition probabilities over the duration of the infectious period. This figure appears in color in the electronic version of this article, and any mention of color refers to that version

uncertainty more accurately captures the variation in the transmission probability, compared to the alternative models. This results in coverage probability close to the reference model, and does not come at the cost of additional bias. From this we can conclude that when estimating transmission probabilities in the presence of diagnostic uncertainty, the diagnostic ILM should be preferred.

5 | APPLICATION - IOWA MUMPS EPIDEMIC

Symptom onset date was used to inform the timing of the transition to the infectious compartment as described in Section 2.2, and individuals with imputed symptom onset dates treated identically to those with the date recorded. To incorporate untested individuals in the analysis, city populations were retrieved from U.S. Census data (U.S. Census Bureau, 2021). An important characteristic of this particular epidemic is that it occurred primarily in college students (Dayan *et al.*, 2008), with close quarters in dorms/classrooms and spring break travel playing an important role in transmission (Polgreen *et al.*, 2010). Although the testing data included the age of each tested individual, the Census data was not categorized by age at the city level.

To account for differences in data availability between tested and untested individuals, different infection kernels were used depending on the test status of susceptible individual i. If individual i was tested ($y_i^T = 1$), the infection kernel is specified as $\kappa(i,j) = \lambda a_{ij} + \delta_{ij}$. Here a_{ij} is an indicator that both the susceptible and infectious individuals are college-age (aged 17–24) and $\delta_{ij} = d_{ij}^{-\rho}$ captures distance-based transmission, where d_{ij} is the Euclidean

distance between the centroids of the cities where individuals i and j reside. Because distance is based on city location, we used $d_0 = 0.01$ for individuals in the same city ($d_{ij} = 0$) as it is smaller than the Euclidean distance between the two closest cities in Iowa, 0.012. For cities with distance above 0.7 units apart (roughly 55 miles), we assume no distance-based transmission ($\delta_{ij} = 0$). If individual i was not tested ($y_i^T = 0$), we use the infection kernel, $\kappa(i, j) = I(d_{ij} = 0)$, which specifies homogeneous mixing between untested susceptibles and tested infectious individuals in the same city.

As in the simulation, infectivity was assumed constant until an exponential decay intervention starting on the date of the CDC dispatch (March 30, 2006). We allow the baseline infectivity to differ based on the testing status, to test the hypothesis that the testing process self-selects a network of individuals more likely to become infectious. We used the sparks term, ϵ, to capture unexplained infectious pressure. The transmission probability can be written as $\pi_{it}^{(SE)} = 1 - \exp[\{-\psi\gamma^{(1-y_i^T)}e^{-\omega_T(t-t^*)_+} \times \sum_j I(Z_{jt}^I = 1)\kappa(i, j)\} - \epsilon]$. The transition from **E** to **I** was assumed constant at 17 days, and the transition from **I** to **R** was modeled using the path-specific formulation of Section 2.2.

The model was fit using the RJ-MCMC method described in Section 3.1. Priors were identical to those used in the simulation, with the additional parameters λ and γ being given Gamma(0.1, 0.1) priors. The sensitivity and specificity priors for the urine test were assigned based on Krause *et al.* (2006) and were Beta(30, 70) and Beta(99, 1), respectively. Details of the MCMC scheme, tuning, and convergence assessment are presented in Web Appendix C.

The primary quantities of interest from this model are the transmission and removal probabilities, as these are more interpretable than the individual parameters. We describe these probabilities using the posterior mean and 95% credible intervals, which are displayed in Figure 3. Web Table 2 summarizes the posterior distribution for all parameters. The transmission probability is impacted by the distance between, testing status, and ages of the susceptible individual and those infectious. To illustrate these effects, we estimate the probability assuming one infectious individual either in the same city or $>$ 55 miles away and for assuming tested individuals are both college-age or not. The removal probability is estimated over the length of an individual's infectious period.

First, we notice that there is not a strong response to the intervention on March 30th, shown by the minimal reduction in $\boldsymbol{\pi}^{(SE)}$ over time, and corresponding with previous modeling of this epidemic (Porter and Oleson, 2013). We also observe a strong age effect, with higher probability of transmission between two college-age individuals living very far apart than two non-college age individuals living in the same city. This provides evidence that a successful intervention targeted at college students would be effective in controlling spread. When distance-based transmission is not included, the credible intervals are narrower reflecting the uncertainty in the estimation of ρ. The model estimates a very low probability of transmission that is unexplained by distance or age. The probability of transition from infectious to removed favors a longer infectious period of between 6 and 10 days. Porter and Oleson (2013) also found that longer infectious periods better described these data. This implies that transmission was occurring more than a week after the onset of symptoms.

6 | DISCUSSION

The diagnostic ILM allows researchers to make inference on important epidemic characteristics in the presence of uncertainty from one or more diagnostic tests. This is an important extension to the stochastic compartmental infectious disease modeling framework, particularly for modeling diseases in which a gold standard test does not exist or when diagnostic accuracy is affected by vaccination status or co-infections. Using a data model on the test results and the Bayesian framework to leverage information about test characteristics and the transmission process, we were able to estimate each individual's latent disease status within the model and accurately capture variability resulting from the testing process. Depending on data availability, the proposed method can be extended to include many different factors, including vaccination status or full network data (e.g., Malik *et al.*, 2014), however, the motivating data did not include individual vaccination records.

This model also illustrates a use for data on negative test results, which are often thrown out and not considered useful for analyses. In the data application, 4,451 of the 6,082 individuals were negative on all tests that they received, and in the diagnostic ILM, incorporation of negative tests are essential for accurate estimation of transmission probabilities. This has important public health implications, stressing the need for recording negative tests as well as positive tests. Moreover, in periods of active surveillance, sampling probabilities could be further incorporated into the diagnostic ILM scheme to provide improved estimates of prevalence in difference communities.

There are a few important limitations to the diagnostic ILM approach. First, computation in the ILM framework is already challenging and time-consuming and the addition of estimation of latent infectiousness status adds to the complexity. The primary cost arises in the repeated calculation of the transmission probabilities at each iteration of the MCMC algorithm, due to the need to recalculate the infection kernel for a large number of susceptible

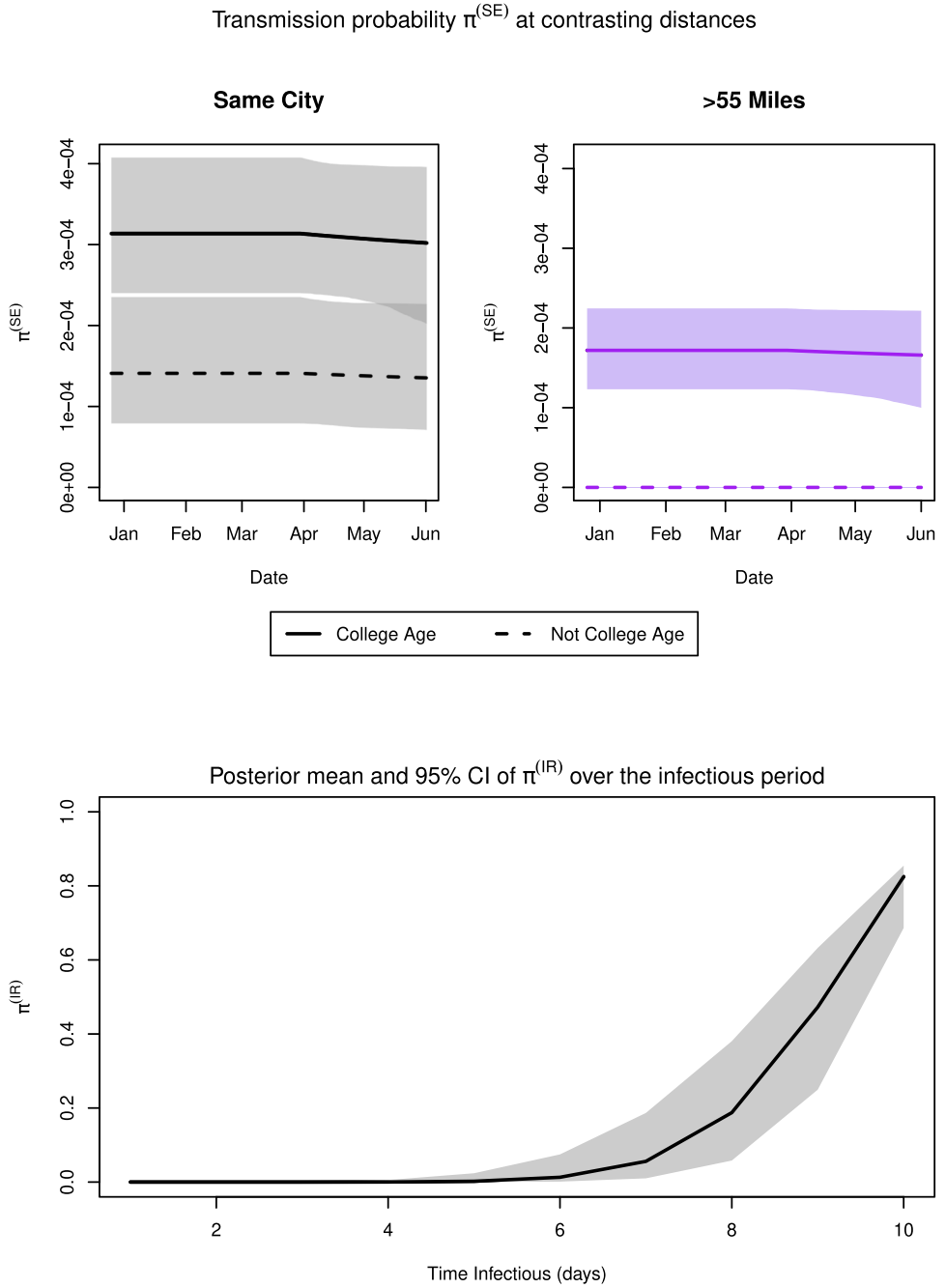

FIGURE 3 Posterior mean and 95% credible intervals for the transmission (top) and removal (bottom) probabilities. This figure appears in color in the electronic version of this article, and any mention of color refers to that version

individuals. Introducing sparsity in the infection kernel, as was done in the data analysis, can reduce computational complexity when proposing an individual's infectious status, as the kernel only needs to be calculated for individuals associated with that one individual. Additionally, several approaches have been taken to ease the computational expense in ILMs, including linearization through Taylor series expansion (Deardon *et al.*, 2010), Gaussian process emulators (Pokharel and Deardon, 2016), and supervised classification techniques (Augusta *et al.*, 2018), and future work can use these techniques to ease the computational challenges of the diagnostic ILM.

The modeling of untested individuals is an important consideration in the diagnostic ILM. In our analyses, we allow a constant sparks term in the transmission probability to capture this infectious pressure. Recent studies of mumps transmission in college students found no evidence of viral shedding in asymptomatic students (Bonwitt *et al.*, 2017), so we believe this assumption is reasonable. However, if the diagnostic ILM was used to

model a disease with a large amount of asymptomatic transmission (e.g., COVID-19), careful consideration of the modeling approach to the sparks term should be taken and contact-tracing data would be imperative for informing this process. In more complex settings, the sparks term can be extended to incorporate individual characteristics or depend on prevalence over time, as in Deardon *et al.* (2010).

Another limitation is that the current model assumes that the sensitivity and specificity of the tests are constant over the period of the infection and are identical for each individual. For mumps, Polgreenet *et al.* (2008) found that the probability of a positive test decreases monotonically after symptom onset. There is also evidence that the sensitivities of the diagnostics change after the onset of symptoms (Sanz *et al.*, 2006; Rota *et al.*, 2013). Additionally, the model assumes conditional independence between the diagnostics. We believe this is a reasonable assumption for our data as exploratory analyses did not indicate strong conditional dependence and this assumption has been used in previous analyses of these three diagnostics (Hatchette *et al.*, 2009). It would be possible to incorporate conditional dependence using methods such as Dendukuri and Joseph (2001) or Georgiadis *et al.* (2003), however, these would need to be extended to allow for incomplete testing results.

In the simulation study, we observed the diagnostic ILM to have the highest amount of bias in estimating the intervention parameter. Further investigation (Web Appendix B) revealed the diagnostic ILM to overestimate the intervention effect, particularly for simulated epidemics with low prevalence and infectious individuals with false negative tests post-intervention. If the primary interest of the analysis lies in estimating the intervention parameter and the epidemic is small, it may be preferable to use standard approaches using any positive test to define the infectious class. However, our simulations showed this approach results in an underestimation of the transmission rate over the course of the epidemic.

Despite these limitations, we have demonstrated that the diagnostic ILM is advantageous in estimating transmission probabilities compared to alternative approaches researchers might attempt. We have also presented a novel approach to estimation of the infectious period within the ILM framework by extending the path-specific modeling approach to data on the individual level. Both developments are demonstrated on a real mumps epidemic, illustrating the types of inference that can be made from the diagnostic ILM. The addition of uncertainty with regards to test results and the length of the infectious period is a more realistic reflection of the disease process and can aid in answering important public health questions.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the Fogarty International Center of the National Institutes of Health under Award Number R01TW010500. The content is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health. We would also like to thank Mike Pentella and the State Hygienic Lab at the University of Iowa for supplying the data. We thank the associate editor and two anonymous reviewers for their insightful comments on previous versions of the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings in this paper were obtained from the University of Iowa State Hygienic Laboratory (SHL). Restrictions apply to the availability of these data, which were used under license in this paper. Data are available on request from the corresponding author with the permission of the SHL. The data are not publicly available due to privacy restrictions. University of Iowa SHL contact information - Address: 2490 Crosspark Road Coralville, IA 52241, Web: http://www.shl.uiowa.edu/contact.

ORCID

Caitlin Ward https://orcid.org/0000-0002-0806-0222

REFERENCES

Akhtar, S. and Carpenter, T.E. (2013) Stochastic modelling of intra-household transmission of hepatitis C virus: evidence for substantial non-sexual infection. *Journal of Infection*, 66, 179–183.

Augusta, C., Deardon, R. and Taylor, G. (2018) Deep learning for supervised classification of spatial epidemics. *Spatial and Spatio-Temporal Epidemiology*, 29, 187–198.

Boddicker, J.D., Rota, P.A., Kreman, T., Wangeman, A., Lowe, L., Hummel, K.B., et al. (2007) Real-time reverse transcription-PCR assay for detection of mumps virus RNA in clinical specimens. *Journal of Clinical Microbiology*, 45, 2902–2908.

Bonwitt, J., Kawakami, V., Wharton, A., Burke, R.M., Murthy, N., Lee, A., et al. (2017) Notes from the field: absence of asymptomatic mumps virus shedding among vaccinated college students during a mumps outbreak–Washington, February–June 2017. *MMWR. Morbidity and Mortality Weekly Report*, 66, 1307.

Brockhoff, H.J., Mollema, L., Sonder, G.J., Postema, C.A., van Binnendijk, R.S., Kohl, R.H., et al. (2010) Mumps outbreak in a highly vaccinated student population, The Netherlands, 2004. *Vaccine*, 28, 2932–2936.

CDC (2006) Update: multistate outbreak of mumps – United States, January 1–May 2, 2006. *Morbidity and Mortality Weekly Report*, pp. 559–563.

CDC (2019) Specimen collection, storage, and shipment. Atlanta, GA: Centers for Disease Control and Prevention.

Chan, J.F., Sridhar, S., Yip, C.C., Lau, S.K. and Woo, P.C. (2017) The role of laboratory diagnostics in emerging viral infections: the example of the Middle East Respiratory Syndrome epidemic. *Journal of Microbiology*, 55, 172–182.

Chen, S., Chang, C., Jou, L. and Liao, C. (2007) Modelling vaccination programmes against measles in Taiwan. *Epidemiology & Infection*, 135, 775–786.

Collins, J. and Huynh, M. (2014) Estimation of diagnostic test accuracy without full verification: a review of latent class methods. *Statistics in Medicine*, 33, 4141–4169.

Dayan, G.H., Quinlisk, M.P., Parker, A.A., Barskey, A.E., Harris, M.L., Schwartz, J. M.H., et al. (2008) Recent resurgence of mumps in the United States. *New England Journal of Medicine*, 358, 1580–1589.

Deardon, R., Brooks, S.P., Grenfell, B.T., Keeling, M.J., Tildesley, M.J., Savill, N.J., et al. (2010) Inference for individual level models of infectious diseases in large populations. *Statistica Sinica*, 20, 239–261.

Dendukuri, N. and Joseph, L. (2001) Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*, 57, 158–167.

Evans, C., Medley, G., Creasey, S. and Green, L. (2010) A stochastic mathematical model of the within-herd transmission dynamics of porcine reproductive and respiratory syndrome virus (PRRSV): fade-out and persistence. *Preventive Veterinary Medicine*, 93, 248–257.

Gastwirth, J.L., Johnson, W.O. and Reneau, D.M. (1991) Bayesian analysis of screening data: application to aids in blood donors. *Canadian Journal of Statistics*, 19, 135–150.

Gelman, A., Rubin, D.B., et al. (1992) Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–472.

Georgiadis, M.P., Johnson, W.O., Gardner, I.A. and Singh, R. (2003) Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52, 63–76.

Green, P.J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82, 711–732.

Hamborsky, J., Kroger, A. and Wolfe, C. (2015) *Mumps*. Centers for Disease Control and Prevention.

Hatchette, T., Davidson, R., Clay, S., Pettipas, J., Leblanc, J., Sarwal, S., Smieja, M. and Forward, K. (2009) Laboratory diagnosis of mumps in a partially immunized population: The Nova Scotia experience. *Can J Infect Dis Med Microbiol*, 20, e157–62.

Heesterbeek, H., Anderson, R.M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., et al. (2015) Modeling infectious disease dynamics in the complex landscape of global health. *Science*, 347, aaa4339.

Hui, S.L. and Walter, S.D. (1980) Estimating the error rates of diagnostic tests. *Biometrics*, 36, 167–171.

Krause, C.H., Eastick, K. and Ogilvie, M.M. (2006) Real-time PCR for mumps diagnosis on clinical specimens–comparison with results of conventional methods of virus detection and nested PCR. *Journal of Clinical Virology*, 37, 184–189.

Malik, R., Deardon, R., Kwong, G.P. and Cowling, B.J. (2014) Individual-level modeling of the spread of influenza within households. *Journal of Applied Statistics*, 41, 1578–1592.

Marin, M., Quinlisk, P., Shimabukuro, T., Sawhney, C., Brown, C. and LeBaron, C.W. (2008) Mumps vaccination coverage and vaccine effectiveness in a large outbreak among college students – Iowa, 2006. *Vaccine*, 26, 3601–3607.

Peltola, H., Kulkarni, P.S., Kapre, S.V., Paunio, M., Jadhav, S.S. and Dhere, R.M. (2007) Mumps outbreaks in Canada and the United States: time for new thinking on mumps vaccines. *Clinical Infectious Diseases*, 45, 459–466.

Pokharel, G. and Deardon, R. (2016) Gaussian process emulators for spatial individual-level models of infectious disease. *Canadian Journal of Statistics*, 44, 480–501.

Polgreen, P.M., Bohnett, L.C., Cavanaugh, J.E., Gingerich, S.B., Desjardin, L.E., Harris, M.L., et al. (2008) The duration of mumps virus shedding after the onset of symptoms. *Clinical Infectious Diseases*, 46, 1447–1449.

Polgreen, P.M., Bohnett, L.C., Yang, M., Pentella, M.A. and Cavanaugh, J.E. (2010) A spatial analysis of the spread of mumps: the importance of college students and their spring-break-associated travel. *Epidemiology and Infection*, 138, 434–441.

Porter, A.T. and Oleson, J.J. (2013) A path-specific SEIR model for use with general latent and infectious time distributions. *Biometrics*, 69, 101–108.

Rota, J.S., Rosen, J.B., Doll, M.K., McNall, R.J., McGrew, M., Williams, N., et al. (2013) Comparison of the sensitivity of laboratory diagnostic methods from a well-characterized outbreak of mumps in New York City in 2009. *Clinical and Vaccine Immunology*, 20, 391–396.

Sanz, J.C., Mosquera Mdel, M., Echevarria, J.E., Fernandez, M., Herranz, N., Palacios, G., et al. (2006) Sensitivity and specificity of immunoglobulin G titer for the diagnosis of mumps virus in infected patients depending on vaccination status. *APMIS*, 114, 788–94.

Trotz-Williams, L., Mercer, N., Paphitis, K., Walters, J., Wallace, D., Kristjanson, E., et al. (2017) Challenges in interpretation of diagnostic test results in a mumps outbreak in a highly vaccinated population. *Clinical and Vaccine Immunology*, 24, e00542–16.

U.S. Census Bureau (2021) City and town population totals: 2010-2019.

SUPPORTING INFORMATION

Web Appendices, Tables, and Figures referenced in Sections 3.1, 4, and 5 are available with this paper at the Biometrics website on Wiley Online Library. The R code for the simulation study is available at the Biometrics website on Wiley Online Library.

How to cite this article: Ward, C., Brown, G.D., Oleson, J.J. An individual level infectious disease model in the presence of uncertainty from multiple, imperfect diagnostic tests. *Biometrics*. 2021;1–11. https://doi.org/10.1111/biom.13579