Risk Analysis, Vol. 21, No. 6, 2001

Poliovirus Surveillance by Examining Sewage Water Specimens: Studies on Detection Probability Using Simulation Models

J. Ranta,[1]* T. Hovi,[2] and E. Arjas[1]

Efficiency of environmental surveillance of poliovirus circulation was studied using simulation models. First, three transmission models were defined for describing different scenarios of poliovirus infections in a large unstructured population. Second, environmental factors, such as the total volume of the sewage network and losses of viruses, were modeled for computing the virus output at the sewage sampling site. Third, the effect of sampling and laboratory procedures was accounted for in the probability of detection, given the amount of polioviruses in a specimen. The simulation model can be used for theoretical assessments of the likely efficiency of environmental surveillance, compared with acute flaccid paralysis (AFP) surveillance. Under reasonable assumptions in a vaccinated population, the AFP surveillance can be outperformed if the poliovirus outbreak is not large. However, this depends on the assumed case-to-infection ratio and on the sampling frequency of the sewage water specimens. Increasing the latter will lead to a higher detection probability, which will further enhance the method based on environmental surveillance.

KEY WORDS: Polio; surveillance; detection probability; simulation models; environmental sampling

1. INTRODUCTION

The program for global eradication of poliomyelitis coordinated by the World Health Organization is progressing well, and, before the end of the first decade of the new millennium, we may realistically consider the option of terminating polio immunizations.[1] Before this can be accomplished, however, there must be a high degree of certainty that there is no remaining circulation of wild polioviruses anywhere in the world. The proportion of people who develop typical paralytic disease form upon poliovirus infection is thought to be between 1% and 0.1% in unvaccinated populations, and may be much less in partially immune groups of people.[2,3] A low-level circulation of poliovirus in human populations may remain undetected if surveillance of acute flaccid paralysis (AFP) is not functioning properly.[4] Population-based screening for carriers of poliovirus is inefficient and expensive, unless it can be targeted to a relatively small group of suspected virus excretors. Therefore, it has been suggested that viruses could be sought from the environment and especially from sewage waters. The rationale behind this suggestion is the fact that every infected individual, whether symptomatic or not, will excrete large amounts of poliovirus into the feces, and thus into the environment, for several weeks. Polioviruses can persist in the environment for variable times, long enough for its detection from the outlet of a sewage network to

[1] Rolf Nevanlinna Institute, University of Helsinki, Finland.
[2] Enterovirus Laboratory, Department of Virology, National Public Health Institute, Helsinki, Finland.
*Address correspondence to J. Ranta, Rolf Nevanlinna Institute, P.O. Box 4, FIN-00014, University of Helsinki, Finland; Jukka.Ranta@RNI.Helsinki.FI.

0272-4332/01/1200-1087$16.00/1

be used as a surrogate of poliovirus-infected individuals in the corresponding region.[5–7]

Environmental surveillance can be efficiently implemented in regions with fully functioning sewage processing networks. There are several examples of wild poliovirus detection in the environment in the absence of reported AFP cases or suspected poliomyelitis.[2,8] In spite of this, no widely accepted strategies for sampling exist, and the methods used for virus detection in the environmental specimens are not standardized. Factors that influence the final population sensitivity of the system, that is, detection of poliovirus circulation in the target population by examining environmental specimens, can be divided into three groups. The first group consists of epidemiologic parameters such as the number of infected individuals and the amount and duration of virus excretion. The second group of factors can be described as environmental and they determine the remaining detectable fraction of the initial input of virus in the compartment under study. The final group, and the only one in which the values can be altered, comprises sampling frequency, sample size, and sensitivity of the laboratory procedures applied.

In the following article, mathematical models are presented for each of the three groups of factors, and examples are given on how the models can be used to quantify the environmental sensitivity under a selected transmission scenario, sampling scheme, and laboratory techniques.

2. METHODS

2.1. Endemic and Epidemic Transmission Scenarios

We formulate three different dynamic models[9–12] of poliovirus circulation in a large population. The models describe the number of carriers of poliovirus as a (random) function $I(t)$ of time. In a stationary model, poliovirus infections occur in a human population with a constant incidence rate $\lambda(t) = \lambda_c$. This can be taken as a simplification of endemic low-rate transmission of poliovirus in the population, or repeated introduction of poliovirus into the population from external sources of infection, but without secondary infections. In this scenario, the number of infections up to time t is then Poisson distributed with mean $\lambda_c t$. Moreover, if the duration of each infection is assumed independent and exponential with mean $1/\mu$, then the incidence rate of ending infections in a population of $I(t)$ carriers is $\mu(t) = \mu I(t)$. In equilibrium, the expected prevalence is $E[I(t)] = \lambda_c/\mu$ for any time t.[13,14] No matter what the initial value $I(0)$ is, this process always reaches its equilibrium. The population size of susceptibles is assumed to be sufficiently large compared with the number of infected individuals, so that there is no limiting effect on the epidemic process. This is an admissible assumption here because in the scenarios we are interested in, the number of infected carriers of poliovirus is relatively low.

The number of infected individuals $I(t)$ is now easy to simulate (see the Appendix) as a step function of time. In order to define some endemic level of carriers $E[I(t)]$, one must choose suitable values for λ_c and μ. In practice, we have some knowledge of the duration of poliovirus infections, which provides us with a rough estimate of μ. This constitutes our first transmission scenario (Model 1):

$$\lambda(t) = \lambda_c, \quad \mu(t) = \mu I(t), \quad I(0) = I_0, \qquad (1)$$

where the incidence rates are given default values $\mu = 1/21$ and $\lambda_c = 100/365$. This would correspond to 3 weeks expected duration of the infectious period and 100 expected new infections per year. The total excretion period[15] is variable and usually much longer than 3 weeks, but the concentration of the virus in stool decreases after 2 weeks. Therefore, exponential distribution with mean equal to 3 weeks seems to be a plausible model. The stationary prevalence is then $\lambda_c/\mu = 5.8$ (default) and the initial value is chosen accordingly as $I(0) = 6$.

A nonstationary model can be introduced in a similar fashion, by replacing the constant incidence rate λ_c by $\lambda I(t)$. Now the incidence rate is proportional to the current number of infected carriers of poliovirus. This model could be used to represent the early stage of an outbreak in a largely susceptible population. The only possible stationary states for the number of carriers are then 0 and "infinity." The expected prevalence follows an exponential function $E[I(t)] = I_0 \exp[(\lambda - \mu)t]$.[13,14] Each infected carrier causes a number of secondary infections with mean $\mathcal{R}_0 = \lambda/\mu$. If this basic reproduction number is less than 1, the epidemic eventually becomes extinct with probability 1. This gives the second transmission scenario (Model 2):

$$\lambda(t) = \lambda I(t), \quad \mu(t) = \mu I(t), \quad I(0) = I_0, \mathcal{R}_0 < 1. \qquad (2)$$

The chosen default values here are $\mu = 1/21, \lambda = 1/23$, and $I(0) = 1$. With these parameters, there is approximately 0.024 (± 0.0013, Monte Carlo error, $10 \times 10{,}000$ simulations) probability that an outbreak survives longer than 365 days. Given that it survives, the expected prevalence $E[I(365) \mid I(365) > 0]$ is approxi-

mately 9 (±1, Monte Carlo error, 10 × 10,000 simulations).

If $\mathcal{R}_0$ is larger than 1, there is a positive probability that an outbreak will grow exponentially, without ever becoming extinct. This is the third transmission scenario (Model 3):

$$\lambda(t) = \lambda I(t),\ \mu(t) = \mu I(t),\ I(0) = I_0,\ \mathcal{R}_0 > 1. \quad (3)$$

Here the default values are $\mu = 1/21$, $\lambda = 1/19$, and $I(0) = 1$. With these parameters, there is approximately 0.11 (±0.0064, Monte Carlo error, 10 × 10,000 simulations) probability that an outbreak survives at least 365 days. Given that it survives, the expected prevalence $E[I(365) \mid I(365) > 0]$ is approximately 57 (±3, Monte Carlo error, 10 × 10,000 simulations). Even though $\mathcal{R}_0$ is larger than 1, most of the epidemics die out rather soon.

In the stationary situation (Model 1), there is no similar formula for $\mathcal{R}_0$, because it was assumed that there are no secondary infections. Neither are there extinction times in Model 1. However, for endemicity, the "effective reproduction number" must be equal to 1 in the sense that the ending infections are expected to be replaced by equally many new infections in the long run.

Both endemic and epidemic scenarios can be used as working hypotheses in the following calculations. Every infected carrier of poliovirus excretes daily a quantity, ρ, of viruses. Given the number of carriers $I(t)$, the total amount of excreted polioviruses per time unit is then $\rho I(t)$, at time t. Here, the time unit is taken to be 1 day. In each case, a population model (1, 2, and 3) of poliovirus production can be given by defining three parameters λ, μ, and ρ, together with some initial condition $I(0)$. For demonstration purposes only, each model is applied using the above default values, with $\rho = 10^7$ $CCID_{50}$ (50% Cell Culture Infectious Dose).[16]

2.2. A Model of Environmental Factors

It is assumed that a large metropolitan area defines our study population and the properties of the local sewage system can be described adequately by the following rough measures. An important quantity is the daily sewage flow, M. In the following, we restrict our considerations to a sewage network in which all peripheral drains eventually converge and bring their contents to a single representative outlet that serves as the sampling site. The sewage system causes some delay and a momentary input load of viruses will appear gradually at the detection site. For the network delay times, one can envisage two extreme models. In one, the network is seen as a tube with constant delay time, whereas in the other it is seen as an instantaneously and completely mixing pool with exponentially distributed delay times. A more realistic model is obtained from a gamma distribution describing a chain of instantaneously mixing pools, each with exponential delay. (The sum of independent exponentially distributed random times has a gamma distribution.[17]) In any case, a distribution function g is specified to account for the inherent delay. When the expected delay is relatively short in comparison to the surveillance time and the temporal dynamics of the epidemic process, the exact functional form of g is not crucial. For simplicity, it may well be assumed that g is exponential

$$g(t|\theta) = \theta \exp(-\theta t),$$

where $1/\theta$ is the expected delay time for an arbitrary polio virus particle. The delay time has recently been investigated in the sewage system of Helsinki. One liter of type 1 poliovirus vaccine was introduced into the sewage and appearance of the virus was monitored at the inlet of a single sewage treatment plant located 20 km downstream in the network and representing a population of 700,000 people. The virus was detected in the first 24-hour sample and was also recovered in samples collected during 3 subsequent days. This can be interpreted as indicating fair mixing of momentarily introduced virus and partial delay of virus transport in a standard city sewerage.[18]

There may also be environment-dependent loss of viruses. The losses may be due to various factors. These can be summarized in terms of total loss percentage, because it is difficult to give reliable estimates of each of the different causes of loss. Some households in the region may not be connected to the sewage network. Use of disposable diapers may divert poliovirus, excreted by babies, from sewage to collections of solid waste. Finally, poliovirus particles may be physically trapped within the sewage network, because they may be chemically or physically inactivated before arrival to the sampling site. The effective initial number of excreted viruses is $V_{\text{in}}(t) = \rho I(t)$, and the total output $V_{\text{out}}(t)$ can now be computed from the input,

$$V_{\text{out}}(t) = \int_0^t \gamma V_{in}(s) g(t - s|\theta) ds,$$

where γ is the fraction of viruses that reach the detection site alive and detectable. This "thinning" factor could also be time dependent if the time spent in the system would play a crucial part in the survival of

the virus. In that case, γ should be replaced by $\gamma(t-s)$ in the above expression. Initial losses caused by other factors would then be described with $\gamma(0) < 1$. The form of a time-dependent loss function is difficult to speculate and the virus survival was therefore described by a constant $0 < \gamma \leqslant 1$. To summarize, the environment is characterized by the mixing properties of the sewage system and the loss percentage. In addition to these, the total daily flow volume M of the system needs to be specified as a diluting factor.

2.3. Modeling the Influence of Sampling Design and Laboratory Procedures

Suitable working hypotheses can be described by the endemic/epidemic and environmental assumptions above. These factors can be quantified but not changed or controlled. In contrast, an investigator can choose how often and how large specimens are taken and what laboratory techniques are used. We first formulate the probability of detection in a single specimen sample, taken at some arbitrary sampling time τ_i. It was previously shown how the output of viruses $V_{\text{out}}(\tau_i)$ is computed, under the given assumptions. The amount of polioviruses, k CCID_{50}, in a specimen sample of volume S is assumed to be distributed according to the Poisson distribution with mean $V_{\text{out}}(\tau_i)S/M$,

$$p_{\tau_i}(k) = \text{Poisson}\left[k; V_{\text{out}}(\tau_i)\frac{S}{M}\right].$$

This assumption can be regarded as a conventional simplification reflecting the lack of more specific knowledge on possible heterogeneity. Other distributional forms would require the specification of several additional parameters that are difficult or impossible to estimate. The total flow of the sewage system M per time unit is in relation to the population size through water consumption per capita. In addition to the domestic waste waters, the total flow volume M includes rainwater, and so forth. If M fluctuates considerably in time it should be modeled as a time-dependent function, but here we assume it to be relatively stable. Given that the amount of viruses in a specimen is k CCID_{50}, the probability that the laboratory technique used gives a positive result, $(+)$, is modeled as

$$p_k(+) = 1 - \exp(-\beta k),$$

where β is a parameter describing the sensitivity of the used technique. This model corresponds to the assumption that the viruses are detected independently, and that there are never false positives. Finally, the overall probability of detection at time τ_i is

$$\pi_i = \sum_{k=1}^{\infty} p_{\tau_i}(k)p_k(+).$$

When samples are taken at many different time points $\tau_1, \ldots, \tau_n$, the cumulative probability of detection before and including sampling time τ_n is

$$P(\text{Detecting polioviruses before time } \tau_n) = 1 - \prod_{i=1}^{n}(1-\pi_i).$$

In decision making, in addition to the sampling times τ_i, the only free variables we can effectively choose are the sample size S and the sensitivity of the used laboratory techniques, expressed via parameter β. The three submodels of (1) endemic/epidemic process (Models 1, 2, and 3), (2) environment, and (3) experimental factors can be represented as a directed graph depicting the result of a single sampling experiment. (Fig. 1).

3. RESULTS

3.1. Comparing Endemic and Epidemic Scenarios

For a particular realization $I(t)$, $t \in (0, T)$ of the endemic/epidemic process, the (momentary) probability of detection, π_i, is different at different sampling times τ_i depending on $V_{\text{out}}(\tau_i)$. This is true for all scenarios. In the sequel, we use the term "conditional probability" to denote a probability that is computed, given a particular realization $I(t)$, or type of realizations. Otherwise, we speak of unconditional probabilities, given only λ, μ and the model, which constitute our "scenario." The latter is perhaps more useful for the present analysis, because we do not want to condition on specific realizations that, in any case, are unknown to us in practice. However, we show examples of conditional probabilities under the condition that the epidemic is still alive at some time point t, that is, $I(t) > 0$.

In a stationary situation (Model 1), the expected prevalence $E[I(t)]$ at any time t is a constant, reflecting a constant external source of infection. Therefore, the expected amount $E[V_{\text{out}}(t)]$ of viruses at sampling site is also a constant. Given only the parameters λ and μ—that is, not conditioning on any particular realization $I(t)$—it follows that the probability of detection π_i is a constant at any sampling time τ_i in a stationary process. Consequently, it is equivalent to collect n samples at different times, or to draw n samples at a same time, or to collect a single n times larger sample. In a series of sampling experiments, the prob-

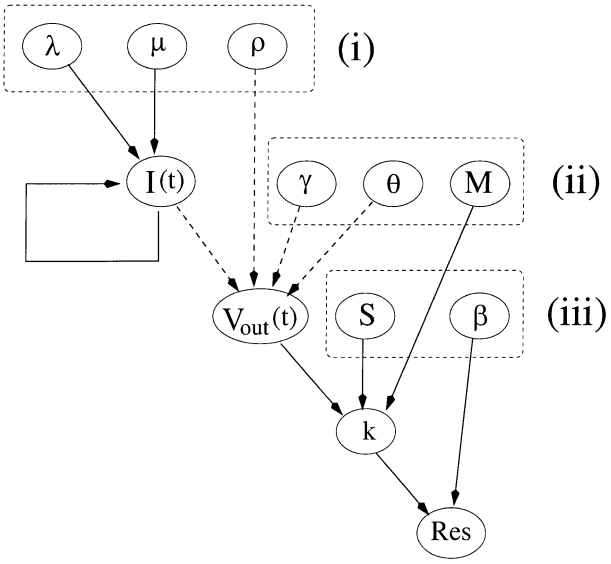

Fig. 1. Graphical representation of the model. Solid arrows denote a stochastic dependence; dashed arrows denote deterministic relations. The three elements of the model, epidemic scenario (i), environment (ii), and survey experiment (iii), are denoted here by grouping the associated parameters in different boxes. λ = incidence rate [total population incidence rate in Models 2 and 3 is $\lambda I(t)$]; $1/\mu$ = expected duration of poliovirus excretion; $I(t)$ = number of polioinfected individuals at time t; ρ = amount of polioviruses excreted by each carrier per day; $(1 - \gamma) \times 100$ = loss percentage of live polioviruses; $1/\theta$ = expected delay (days) caused by the sewage network; $V_{out}(t)$ = amount of polioviruses at sampling site at time t; M = total sewage flow in a day (dilution); S = specimen sample volume; k = number of polioviruses ($CCID_{50}$) in a specimen; β = laboratory parameter controlling specimen sensitivity; Res = laboratory result of a specimen (+/−).

ability of eventually detecting polio increases in time, because at any time τ_i of sampling, there is a positive chance of detection, $\pi_i > 0$. It is of interest to study how fast this probability reaches 1.

The detection problem becomes quite different in the epidemic scenario with $\mathcal{R}_0 = \lambda/\mu < 1$ (Model 2). When $\mathcal{R}_0 < 1$, the epidemic becomes extinct with probability 1. Therefore, the probability of detection in a single sampling experiment π_i decreases to 0 as the sampling time τ_i increases. The conditional cumulative probability of detection reaches 1, given that the virus did not become extinct. The unconditional cumulative detection probability reaches a limit, smaller than 1. When $\mathcal{R}_0 > 1$ (Model 3), there is a positive (not 0) probability that the epidemic grows without a limit. This probability can be very small if $\mathcal{R}_0$ is only a little above 1. The cumulative probability of detection at time t depends on the probability of no extinction before time t.

To summarize, qualitatively, epidemics with any $\mathcal{R}_0$ (Models 2 and 3) may escape our sampling attempts forever with a non-0 probability if there are no persistent external sources of infection, as there are in model 1. In epidemic transmission scenarios, it is of interest to see how high the cumulative probability of detection increases as a function of sampling time, and how fast the cumulative conditional probability increases to 1, given that the epidemic did not die out.

In the following numerical examples (Figs. 2, 3, and 4), we adopt the sampling protocol of one specimen sample every second week. Model parameters {(i): λ, μ, ρ}, {(ii): γ, θ, M}, {(iii): S, β} can be given different values representing different assumptions on the epidemics, the environment, and the quality of laboratory techniques. The default values were:

i. The default values in each of the transmission Models 1, 2, and 3, with $\rho = 10^7$ $CCID_{50}$.
ii. The environmental parameters were chosen as $1/\theta = 3$ days (expected network delay), $1 - \gamma = 0$ (total virus loss), and $M = 310 \times 10^6$ liters (daily sewage flow). These figures are derived from experimental data concerning the sewage network of Helsinki.[18]
iii. The experimental parameters were chosen as $S = 1$ L and $\beta = 4.6$, corresponding to sample size of 1 L and 99% probability of giving a positive result when there is 1 $CCID_{50}$ of viruses in the specimen sample.

3.2. Expected Results of AFP Surveillance

The proposed environmental surveillance scheme may be compared with the alternative procedure of detecting the existence of poliovirus in a population by the notification and investigation of emerging AFP cases. Each polio infection may lead to AFP with some probability, which is thought to be in the range 1/1000–2/100, but is likely to be lower in a vaccinated population.[19] Assume first that polioviruses are always correctly identified from any clinical AFP case whenever poliovirus was the true causative agent. The event times of such notified AFP cases due to polio can be modeled by a thinned infection process with intensity $p\lambda(t)$, where p is the probability of AFP on infection (case-to-infection ratio) and $\lambda(t)$ is the subclinical (latent) infection intensity function according to the transmission model (1, 2, or 3). Heterogeneity in the chances of disease on infection is not accounted for when p is treated as a constant. A reasonable alternative would be to assign a probability density

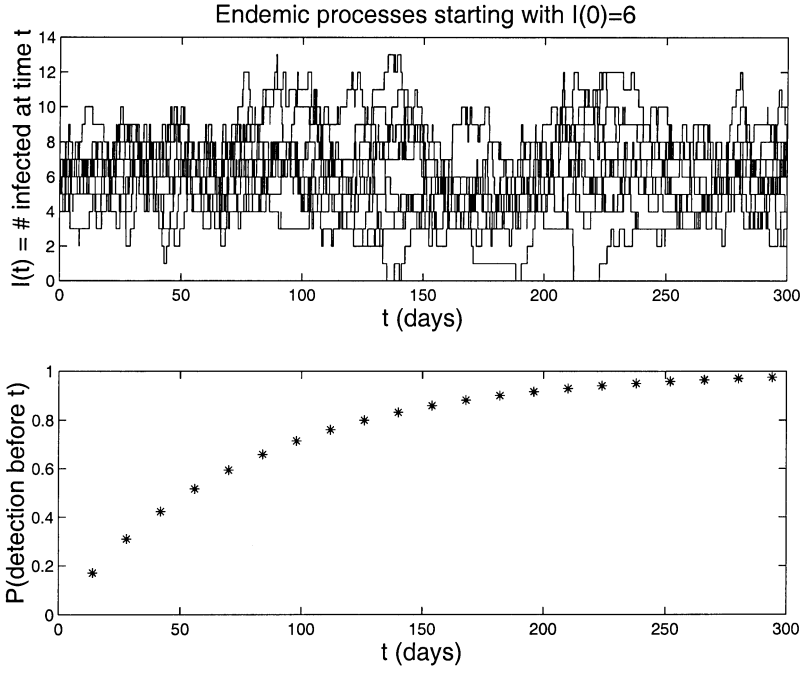

Fig. 2. Some realizations from a stationary scenario (Model 1, $\lambda = 100/365$, $\mu = 1/21$) starting with six polio-infected individuals. The lower figure represents the cumulative probabilities of detection computed at times $\tau_i = 14, 28, \ldots,$ 294 days.

$P(p)$. If polioviruses cannot always be identified from a case of polio disease, then there is some additional thinning, and the parameter p actually represents the fraction of infections leading to AFP that *are* identified as "AFP due to polio." Either way, it follows from a standard result of intensity models that the probability of detecting the first AFP caused by poliovirus before time t is then

$$P\ (\text{Detecting 1st polio AFP case before } t)\ = \\ 1 - \exp\Big[-\int_0^t p\lambda(s)ds\Big].$$

In the endemic scenario (Model 1), the integral inside the exponential is simply $tp\lambda$. Hence, we can compute the above probability as a function of time, and compare this with the cumulative detection probability arising from the sewage water sampling. In Models 2 and 3, the above probability can be computed approximately from n simulations of $I(t)$, by simulating also the AFP occurrences, as

$$P\ (\text{Detecting 1st polio AFP case before } t) \approx \\ \frac{1}{n}\sum_{i=1} Y_i(t),$$

where the indicator function Y is defined for the ith simulated realization as

$$Y_i(t) = \begin{cases} 1, \text{ if the first (simulated) identified polio} \\ \quad \text{disease occurred before } t \\ 0, \text{ otherwise.} \end{cases}$$

Similarly, we can compute conditional detection probabilities under the condition of no extinction. The approximating formula is then defined for those simulated realizations satisfying this condition:

$$P\ (\text{Detecting 1st polio AFP case before } t | I(t) > 0) \approx \\ \frac{\sum_{i=1}^n Y_i(t)U_i(t)}{\sum_{i=1}^n U_i(t)}$$

where

$$U_i(t) = \begin{cases} 1, \text{ if the (simulated) epidemic died out} \\ \quad \text{before } (t) \\ 0, \text{ otherwise.} \end{cases}$$

In Fig. 5 the detection probabilities are computed for each of the three transmission models with $p = 1/1000$ and $p = 1/200$. Eichner and Dietz[9] chose $p = 1/200$ in their simulation studies for a developing country without vaccination. In our example, we aim to describe the situation in a vaccinated population. Therefore, the chances of paralytic poliomyelitis on infection are lower. With $p = 1/1000$, the AFP detection probabilities in Models 2 and 3 remain less than

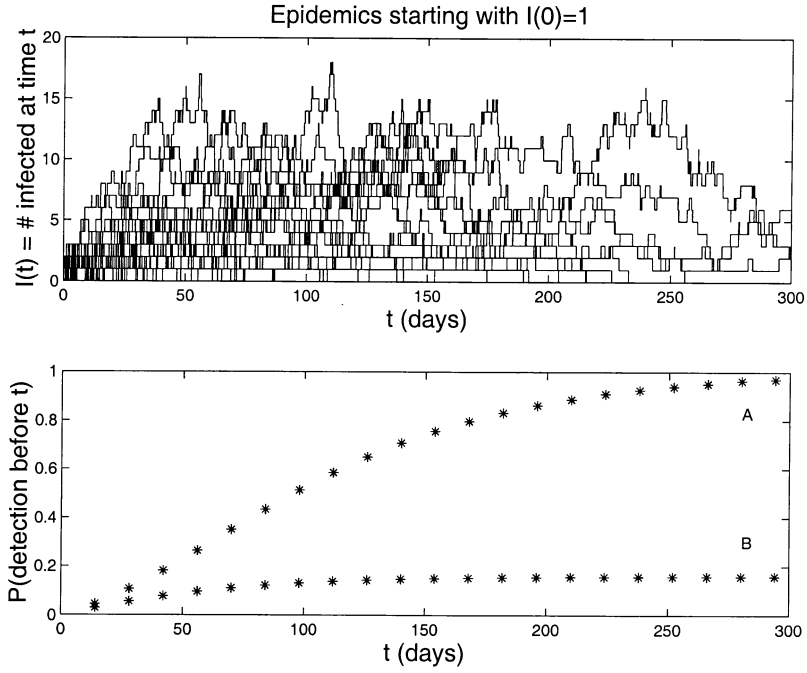

Fig. 3. Some realizations from an epidemic scenario (model 2, $R_0 = 21/23$) starting with one polio-infected individual. The lower figure represents the cumulative probabilities of detection under two assumptions: (A) conditional probability under the event "no extinction in t days," and (B) unconditionally. The probabilities are computed at times $\tau_i = 14, 28, \ldots, 294$ days.

0.04 in 300 days. However, under the condition of no extinction, the detection probabilities will rise to be approximately 0.23 (Model 3) and approximately 0.09 (Model 2) in 300 days. In the endemic scenario (Model 1), it will reach approximately 0.08. These numbers are only suggestive, unless we are confident about the validity of the quantified transmission scenario and the case-to-infection ratio. With $p = 1/200$, the conditional detection probabilities under no extinction rise to approximately 0.65 (Model 3) and 0.4 (Model 2) in 300 days. The dashed lines in Fig. 5 represent Model 2 and are comparable with Fig. 6 de-

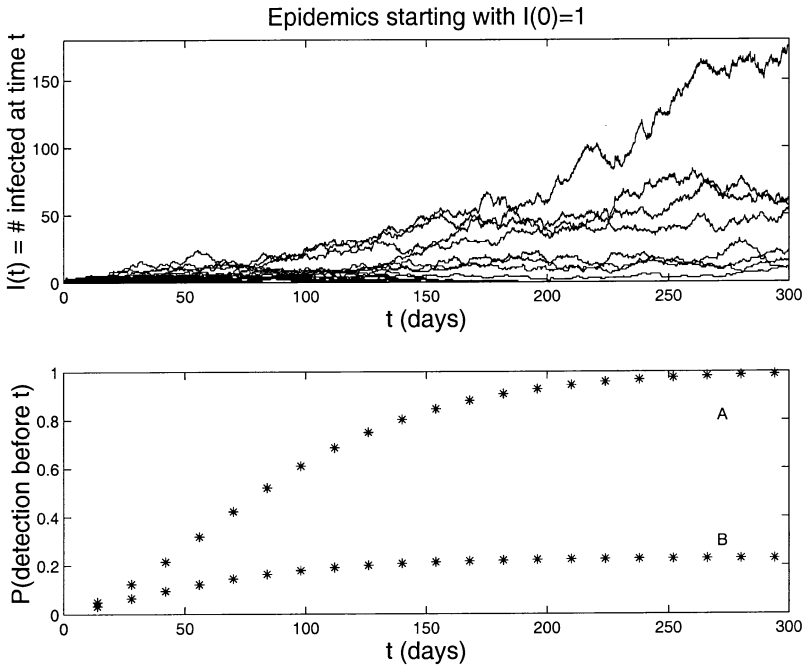

Fig. 4. Some realizations from an epidemic scenario (Model 3, $R_0 = 21/19$) starting with one polio-infected individual. The lower figure represents the cumulative probabilities of detection under two assumptions: (A) conditional probability under the event "no extinction in t days," and (B) unconditionally. The probabilities are computed at times $\tau_i = 14, 28, \ldots, 294$.

rived from the sewage water sampling scheme using Model 2.

3.3. Sensitivity Analysis of Environmental Surveillance

Some aspects of the model assumptions are likely to be more realistic than others. In an optimistic scenario, the quality of laboratory techniques was assumed such that one $CCID_{50}$ of polioviruses in a specimen liter leads to detection with probability $p_1(+) = 0.99$; that is, according to the model, $1 - \exp(-\beta) = 0.99$, giving $\beta = 4.6$. Some other numerical values may be more realistic, however. For example, assuming this specimen sensitivity to be 0.99, 0.5, or 0.1 leads to three different detection probability curves in Fig. 6. These were drawn for the epidemic transmission scenario of Model 2, that is, $\mathcal{R}_0 < 1$. All the other parameters $I(0)$, λ, μ, ρ, γ, θ, M, and S were given the same default values as before.

Similarly, with Model 2, assuming the loss percentage $1 - \gamma$ to be 0%, 20%, and 80%, leads to the curves in Fig. 6. With either low specimen sensitivity (0.1), or large loss percentage (80%), the conditional probabilities of detection rise to be approximately 0.4 and 0.6 in 300 days, respectively. Using AFP surveillance, the conditional probability remains approximately at 0.09 with $p = 1/1000$, but at 0.4 with $p = 1/200$. The comparison is fair in the sense that the underlying transmission model was the same, that is, model 2. For other models, the result is qualitatively the same. Working with unconditional detection probabilities, sewage water sampling is still likely to perform equally or better than AFP surveillance (Figs. 5 and 6). Of course, the result depends on the quality of the environmental surveillance. If all the model parameters are given the worst possible values (as specified above), and sampling is infrequent, then AFP surveillance may perform better. However, it should be noted that we can always increase the environmental sampling frequency to obtain higher detection probabilities. Hence, environmental surveillance can be seen as an active mode of investigation in comparison to passive AFP surveillance in which we are only waiting for the cases. The results indicate that environmental surveillance can be better than AFP surveillance, although this depends on the case-to-infection ratio. If it is higher than $p = 1/200$, AFP surveillance can perform as well or better than environmental surveillance when the environmental loss percentage of viruses is high and/or the laboratory techniques are poor. Nevertheless, the primary aim here was to provide a theoretical framework for quantitative considerations instead of providing parameter values for any particular environment or population.

4. DISCUSSION

We simulated epidemic processes of poliovirus infections according to three simple transmission models. Conditionally, given a transmission scenario,

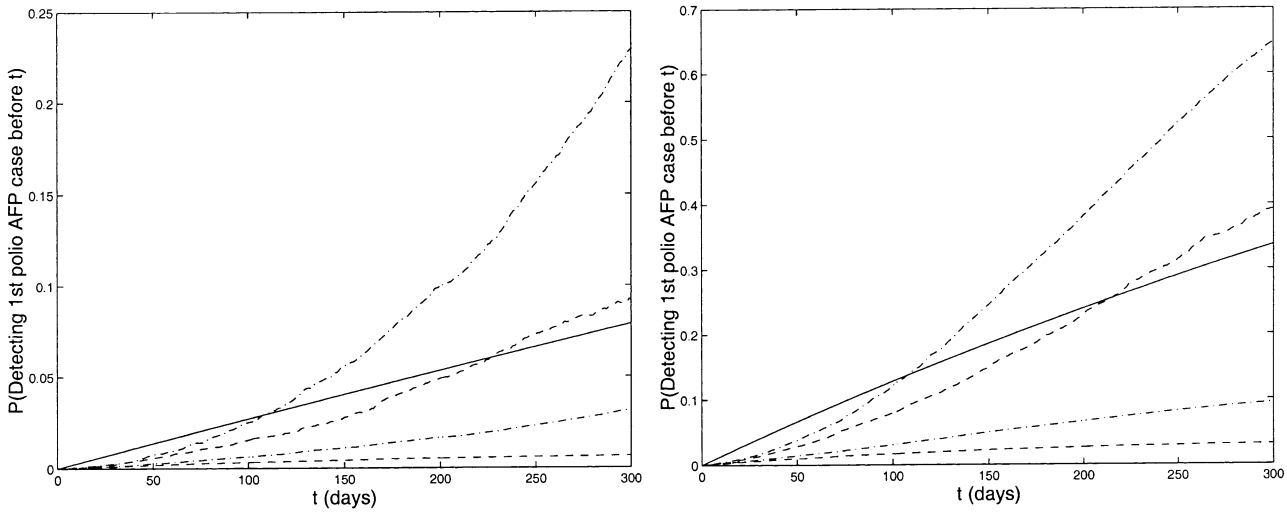

Fig. 5. Detection probabilities of paralytic poliomyelitis (AFP) in three population transmission models: Model 1 (solid line), Model 2 (dash), and Model 3 (dot–dash). The two lower curves (dash and dot–dash) denote unconditional detection probabilities, whereas the two upper curves were computed on the condition of no extinction in t days. The solid curve is computed from the exact formula $1 - \exp(-tp\lambda)$, but the others are based on 50,000 simulations. Left: $p = 1/1000$. Right: $p = 1/200$.

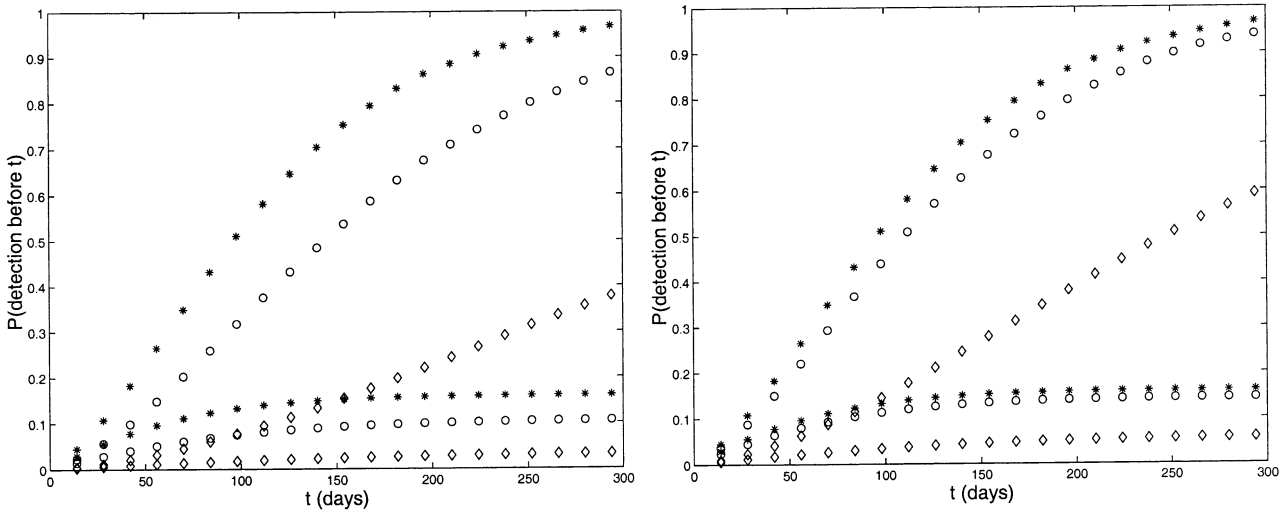

Fig. 6. Left: Cumulative detection probabilities for different specimen sensitivities: $p_1(+) = 0.99$ (asterisk), $p_1(+) = 0.5$ (circle), and $p_1(+) = 0.1$ (diamond). Right: Cumulative detection probabilities for different loss proportions: $1 - \gamma = 0\%$ (asterisk), $1 - \gamma = 20\%$ (circle), and $1 - \gamma = 80\%$ (diamond). In each case, the cumulative detection probabilities are computed unconditionally (lower curve) and by conditioning on the epidemic process not being extinct (upper curve). Model 2: $\lambda = 1/23, \mu = 1/21, I(0) = 1$. The probabilities are computed at times $\tau_i = 14, 28, \ldots, 294$.

the detection probabilities of poliovirus circulation were computed as a function of time. These probabilities depended on the sampling frequency of sewage water specimens, the sample size, and the efficiency of the laboratory techniques. Each of these factors could be quantified within the simulation model framework. This modeling may be helpful for laboratories in considering how to improve the detection sensitivity in environmental surveillance. An approach involving larger or more frequent samples might be more easily achieved than a nominal increase in the sample sensitivity. This model gives a computational method to approximate the expected improvement of sensitivity due to a given change in these parameters. The most important and obvious environmental factors, such as the loss percentage of viruses, delay times, and so forth, can also be quantified within the model.

For a comparison, detection probabilities of AFP cases were computed according to the same transmission models. From this comparison, it appeared that the sewage water sampling scheme could detect polioviruses more efficiently than AFP surveillance. In all transmission scenarios with several parametric assumptions, the sewage sampling outperformed AFP surveillance, assuming that the case-to-infection ratio was less than 1/200 and all AFP cases were promptly diagnosed. This is not, by any means, to suggest that a functioning AFP surveillance should be replaced with environmental surveillance. Our calculations are based on the assumption that a poliovirus detected in an environmental sample is a wild-type poliovirus. As such, this assumption is valid only in populations using the inactivated poliovirus vaccine (IPV), rather than the live virus containing oral poliovirus vaccine (OPV), if standard cell culture isolation is used for virus detection. This is because excess of OPV-derived viruses excreted by the vaccines may, under these conditions, dilute out the minority of wild-type poliovirus if present. However, there are several examples of detection of wild-type poliovirus in the sewage derived from OPV using populations in the absence of reported AFP cases. The problem is partially overcome by the greater fitness of wild-type polioviruses under standard culture conditions, and the culture methods can be modified to further favor the isolation of wild-type viruses.[20] Another limitation of the environmental surveillance is that it can be very difficult to cover the entire population. Furthermore, it is likely that plausible estimates of the environmental factors can be given only for districts in which the sewage system is well controlled and where measurements are reliable. Field experiments may be needed to estimate the loss percentage of viruses, delay times, and other characteristics that are specific to the local environmental conditions. Instead of giving specific point estimates of these parameters, it might be more appropriate to quantify a range of plausible values in the form of a probability distribution describing the uncertainty involved. As a further development of statistical analysis, this would naturally lead to a Bayesian analysis[21]

in which the conditional distribution of disease prevalence were computed under a given a priori scenario, and given the observed specimen sample results.

APPENDIX: SIMULATION IN DIFFERENT TRANSMISSION MODELS

Each of the epidemic models can be simulated by starting with some initial condition $I(0) = I_0 \geq 0$ and the intensities $\lambda(0)$ and $\mu(0)$. The first event time t_1 is then exponentially distributed with parameter $\lambda(0) + \mu(0)$. Random variables from $\exp(\theta)$ distribution are easily drawn by computing $-\log(r)/\theta$, where r is a uniform random variable $r \sim U(0, 1)$. After the event time is drawn from this exponential distribution, the event type is sampled. With probability $\lambda(0)/[\lambda(0) + \mu(0)]$, the first event is "a new infection"; otherwise it is "an ending infection." The step function is updated, $I(t_1) = I(0) \pm 1$, according to the event. After this, the intensity functions $\lambda(t)$ and $\mu(t)$ are updated, because they (in general) depend on the prevalence $I(t)$. Hence, the procedure is repeated until the event time t_n is larger than the end point of the required time interval.

REFERENCES

1. World Health Organization. (1999). *Fourth informal consultation of the polio laboratory network, Geneva, 1–2 October 1998. WHO, Geneva, Switzerland, WHO/V&B99.05*, 11–12.
2. Hovi, T., Cantell, K., Huovilainen, A., Kinnunen, E., Kuronen, T., Lapinleimu, K., Pöyry, T., Roivainen, M., Salama, N., Stenvik, M., Silander, A., Thodén, C. J., Salminen, S., & Weckström, P. (1986). Outbreak of poliomyelitis in Finland: Widespread circulation of antigenically altered poliovirus type 3 in a vaccinated population. *The Lancet, 1*, 1427–1432.
3. Conyn-van Spaendonck, M. A. E., Oosvogel, P. M., van Loon, A. M., van Wijngaarden, J. K., & Kromhout, D. (1996). Circulation of poliovirus during the poliomyelitis outbreak in the Netherlands in 1992–1993. *American Journal of Epidemiology, 143*, 929–935.
4. Birmingham, M. E., Linkins, L. W., Hull, B. P., & Hull, H. F. (1997). Poliomyelitis surveillance: The compass for eradication. *Journal of Infectious diseases, 175*(Suppl 1), S146–S150.
5. Pöyry, T., Stenvik, M., & Hovi, T. (1988). Viruses in sewage waters during and after a poliomyelitis outbreak and subsequent nationwide oral poliovirus vaccination campaign in Finland. *Applied Environmental Microbiology, 54*(2), 371–374.
6. Böttiger, M., & Herrstrom, E. (1992). Isolation of polioviruses from sewage and their characteristics: experience over two decades in Sweden. *Scandinavian Journal of Infectious Diseases, 24*, 151–155.
7. van der Avoort, H. G., Reimerink, J. H., Ras, A., Mulders, M. N., & van Loon, A. M. (1995). Isolation of epidemic poliovirus from sewage during the 1992–3 type outbreak in The Netherlands. *Epidemiology and Infection, 114*, 481–491.
8. Manor, Y., Handsher, R., Halmut, T., Neuman, M., Bobrov, A., Rudich, H., Vonsover, A., Shulman, L., Kew, O., & Mendelson, E. (1999). Detection of poliovirus circulation by environmental surveillance in the absence of clinical cases in Israel and the Palestinian authority. *Journal of Clinical Microbiology, 37*, 1670–1675.
9. Eichner, M., & Dietz, K. (1996). Eradication of poliomyelitis: when can one be sure that polio virus transmission has been terminated? *American Journal of Epidemiology, 143*, 816–822.
10. Keen, R. E., & Spain, J. D. (1992). *Computer simulation in biology.* New York: Wiley-Liss.
11. Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases.* London: Charles Griffin & Company.
12. Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans.* New York: Oxford University Press.
13. Grimmett, G. R., & Stirzaker, D. R. (1994). *Probability and random processes, problems and solutions.* Oxford, UK: Oxford Science Publications, Oxford University Press.
14. Karlin, S., & Taylor, H. M. (1975). *A first course in stochastic processes.* San Diego, CA: Academic Press.
15. Alexander, J. P., Jr., Gary, H. E., Jr., & Pallansch, M. A. (1997). Duration of poliovirus excretion and its implications for acute flaccid paralysis surveillance: A review of the literature. *Journal of Infectious Diseases, 175*(Suppl. 1), S176–S182.
16. Piirainen, L., Stenvik, M., Roivainen, M., Eskola, J., Beuvery, E. C., & Hovi, T. (1999). Randomized, controlled trial with the trypsin-modified inactivated poliovirus vaccine: Assessment of intestinal immunity with live challenge virus. *Vaccine, 17*, 1084–1090.
17. Lindgren, B. W. (1976). *Statistical theory.* New York: Collier Macmillan.
18. Hovi, T., Stenvik, M., Salonen, H., & Kangas, A. (2001). Poliovirus surveillance by examining sewage specimens. Quantitative recovery of virus after introduction into sewage at remote upstream location. *Epidemiology and Infection, 127*, 101–106.
19. Gary, H. E., Jr., Sanders, R., & Pallansch, M. A. (1997). A theoretical framework for evaluating the sensitivity of surveillance for detecting wild poliovirus. II. Factors affecting detection sensitivity in a population with circulating wild poliovirus. *Journal of Infectious diseases, 175*(Suppl. 1), S141–S145.
20. Manor, Y., Handsher, R., Halmut, T., Neuman, M., Abramovitz, B., Mates, A., & Mendelson, E. (1999). A double-selective tissue culture system for isolation of wild-type poliovirus from sewage applied in a long-term environmental surveillance. *Applied and Environmental Microbiology, 65*, 1794–1797.
21. Spiegelhalter, D. J., Myles, J. P., Jones, D. R., & Abrams, K. R. (1999). An introduction to Bayesian methods in health technology assessment. *British Medical Journal, 319*, 508–512.