Supporting Text

Effects of Prior Distribution of R_0

We examined the effects of the prior on R_0 in three ways. First, we recalculated the posterior distribution under three additional prior scenarios on R_0: (*i*) a grave scenario, with the prior median of R_0 set to a relatively high value of 15.4; (*ii*) an optimistic scenario, with the prior median of R_0 set to a relatively low value of 1.8; and (*iii*) a pessimistic scenario with the prior median of R_0 set to a moderately high value of 7 (Table 2). The median values of the optimistic and pessimistic prior distributions reflect the low and high values of R_0 assumed in a previous model of smallpox (1). An important observation is that the curvature of the prior tails is important for the length of the posterior tails and, consequently, for the reported credible intervals (CIs). Although the posterior point estimates and CIs consequently differ somewhat among the four sets of prior scenarios, the comparison between the two intervention strategies nevertheless remains essentially unchanged. Mass vaccination is always favored over trace vaccination regardless of the prior distribution on R_0. All other parameters' prior distributions purposefully were made informative based on the available literature, and, thus, we did not explore the sensitivity to those priors.

The changes in the posterior estimates of R_0 as the prior changed were roughly what one would expect (Table 2 and Fig. 4). In particular, there was considerable overlap across estimates of the posterior. For Old World populations, the median value of the estimate shifted only under the pessimistic prior, which was a result of this prior being much narrower than the other priors under consideration (Fig. 4*C*). For New World populations, the median values were similar except under the optimistic prior (Fig. 4*B*), for which the median of the distribution was higher. The shift to larger values under the optimistic prior occurred because the tail of the log-normal prior on R_0 is essentially flat in the region for which the likelihood is also flat. At higher values of R_0, it therefore becomes increasingly difficult to differentiate between likely draws of R_0 under the Markov chain Monte Carlo (MCMC) sampler, as in Fig. 1. The long tail on the posterior reflects this lack of

Fig. 4. Effects of changes in the prior distribution (solid line) on the posterior histogram of R_0 for Old-World (A,C, and E) and New-World (B,D, and F) populations. (A) and (B) are for an optimistic scenario with a median of 1.8, (C) and (D) are for a grave scenario with a median of 15.4, and (E) and (F) are for a pessimistic scenario with a median of $R_0 = 7$. Note changes in scale.

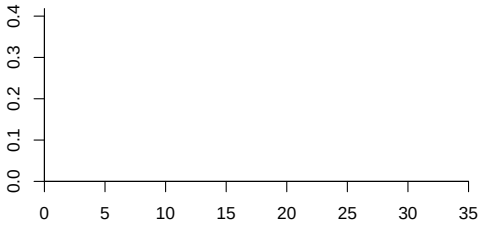


Fig. 5. Difference in the number of fatalities using trace vaccination rather than mass vaccination for Old-World (A and C) and New-World (B and D) populations. In (A) and (B), diagnosis occurs after macales, while in (C) and (D) diagnosis occurs only after pustules appear. All cases assume an optimistic prior distribution on R_0. Note changes in scale.

Fig. 6. As in Fig. 5, but for a grave prior distribution, such that the median on R_0 = 15:4.

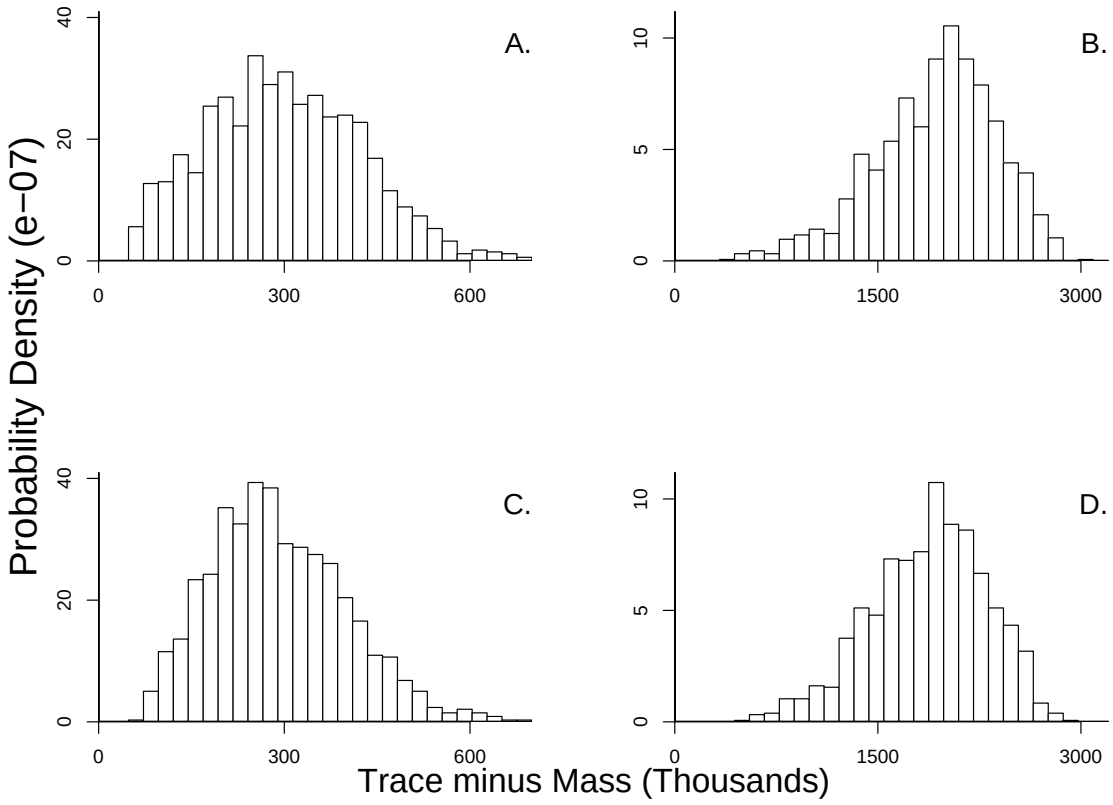


Fig. 7. As in Fig. 5, but for a pessimistic prior distribution, such that the median on $R_0 = 7$.

Table 2. Sensitivity of prior estimates of R_0 (with 95% CI) to the median posterior value of R_0 (with 95% CI).

Estimate	Prior	Posterior
Original – New World	4 (0.7, 23.7)	7.1 (5.0, 15.3)
Old World		17.5 (5.9, 59.5)
Optimistic – New World	1.8 (0.13, 24.2)	7.1 (5.3, 12.2)
Old World		26.9 (9.5, 88.4)
Grave – New World	15.4 (9.4, 25.1)	12.1 (7.5, 21.7)
Old World		18.6 (11.7, 29.2)
Pessimistic – New World	7 (1.9, 26.3)	6.9 (5.3, 12.8)
Old World		19.2 (7.7, 45.5)

Advantages of a Bayesian Approach

The Bayesian approach that we use allows us to incorporate information from as many data sets as possible, including both hospital data on smallpox progression and multiple historical epidemics. Although non-Bayesian methods such as a bootstrap of parameter values (3) may be able to incorporate certain features of this analysis, such methods most likely would fail to account for any potential nonnormality or multimodality of the posterior distribution.

An additional advantage of a Bayesian approach comes from its ability to take true expert opinion into account when establishing prior distributions regarding potentially unknown parameters. For example, we relied on informed expert opinion to center the distribution of mortality rates for Old World populations (4) and, thus, to combine the mortality data coherently with the likelihood. Given these issues, the Bayesian framework represents a particularly promising approach for combining disparate information, expert opinion, and multiple data sets when estimating the parameters of epidemic models.

Fig. 8.

Untraced Individuals –

$$\frac{dS^0}{dt} = I_3 S^0 \quad [c \quad pR_0(t)] \frac{S^0}{N} r_3 I_3 \qquad (4)$$

$$\frac{dI_1^0}{dt} = I_3 S^0 \quad [c \quad pR_0(t)] \frac{I_1^0}{N} + p_{\ 1}(t) \quad r_3 I_3 \quad r$$

have been named by an infected individual enter the queue to be vaccinated. For individuals in the queue to be vaccinated, the equations comprise:

Queued Individuals –

$$\frac{dQ_0}{dt} = [c \quad pR_0(t)] \left(\frac{S^0}{N} r_3 I_3 \quad I_3 Q_0 \quad Q_0 min(1; n=Q)\right) \tag{9}$$

$$\frac{dQ_1}{dt} = I_3 Q_0 + [c \quad pR_0(t)] \frac{I_1^0}{N} + p_{\ 1}(t) \quad r_3 I_3 t$$

Traced but Unsuccessfully Vaccinated Individuals –

$$\frac{dS^1}{dt} = (1 \quad f)(1 \quad {}_0)\ Q_0 min(1; n/Q) \quad S^1 I_3 \tag{14}$$

$$\frac{dI_1^1}{dt} = \quad S^1 I_3 + (1 \quad f)(1 \quad {}_1)\ Q_1 min(1; n/Q) \quad r_1 I_1^1 \tag{15}$$

$$\frac{dI_2^1}{dt} = r_1 I_1^1 + (1 \quad f)\ Q_2 min(1; n/Q) \quad r_2 I_2^1 \tag{16}$$

$$\frac{dI_3^1}{dt} = r_2 I_2^1 + (1 \quad f)(1 \quad h)\ Q_3 min(1; n/Q) + \quad H \quad r_3 I_3^1 \tag{17}$$

$$\frac{dI_4^1}{dt} = r_3(I_3^1 + Q_3 + H) \quad r_4 I_4^1 \tag{18}$$

Here, $_j$ is the vaccine efficacy for stage j, and a superscript of 1 denotes individuals who have been traced but unsuccessfully vaccinated, as previously noted.

The final two equations track the number of recovered-and-immune individuals and the number of dead. These equations take the following form.

Recovered-and-Immune and Dead Individuals –

$$\frac{dZ}{dt} = (1 \quad f)(\ {}_0 Q_0 + \ {}_1 Q_1)\ \ min(1; n/Q) + (1 \quad) r_4(I_4^0 + I_4^1) \tag{19}$$

$$\frac{dD}{dt} = f \quad Q\, min(1; n/Q) + \ r_4(I_4^0 + I_4^1) \tag{20}$$

Here, Z and D represent the number of immune/recovered individuals and the number of dead, respectively. is the death rate for smallpox.

In our analyses, all values derived from the multivariate draws of the posterior are placed into the above ODEs and used to determine the number that die under trace versus mass vaccination (see

Variability in the Parameters of the Vaccination Model

All vaccination parameters were set at Kaplan *et al.*'s (5) baseline rates except for p, the fraction of contacts named, which was conservatively set to 1.0. Also, we assumed that individuals in the early stages of the disease, specifically the first 27% of the latency stage prior to the outbreak of the fever, can be successfully vaccinated. Another difference is that Kaplan *et al*. (5) assume that the prodromal or fever period is infectious; because informed opinion suggests that it is not (6, 7), we instead assume that the prodromal period is noninfectious. Individuals in our model therefore are neither diagnosed as having smallpox nor automatically quarantined until they start exhibiting obvious signs of the disease. Because the early phases of smallpox can be easily confused with the symptoms of other infectious diseases (7), and because few currently practicing physicians have ever diagnosed smallpox, it is unclear how rapidly infected individuals could be diagnosed. We therefore varied the percentage of the infectious period that an infected individual could circulate between 23% and 38% of the overall infectious period. These values correspond to diagnosis at the macales stage, at which flat, discolored patches appear on the skin, and the pustules phase, at which blisters form (7).

Because some have argued that the vaccination rates in Kaplan *et al*. (5) are overly pessimistic, it was also important to consider more optimistic scenarios. We therefore considered two more optimistic vaccination scenarios, in addition to the scenario used in the main text, and reran the queuing model (5) for the case in which we allow for uncertainty in all parameters. Under a moderately optimistic scenario, we increased the number of traced social contacts from 50 to 300, and we increased the number of vaccinators from 5,000 to 10,000. As Fig. 9 shows, under this scenario, the median difference between the trace and mass vaccination strategies when using R_0 estimates based on Old World populations is reduced to 17,000, but there is still a substantial probability that the difference will be higher than 25,000. Similarly, for New World populations, the median difference between the two strategies is reduced to 181,000, but there is still a substantial probability that

the difference will be higher than 1 million. Under a wildly optimistic scenario, we increased the number of traced social contacts from 50 to 1.25 million, and the number of vaccinators from 5,000 to 125 million. As Fig. 10 shows, under this scenario, trace vaccination is more effective than mass vaccination when we use R_0 estimates based on Old World populations, because trace vaccination is virtually identical to mass vaccination. Nevertheless, when we use R_0 estimates based on New World populations, there is still a strong possibility that trace vaccination will produce 1 million or more additional deaths relative to mass vaccination, and uncertainty is still very high.

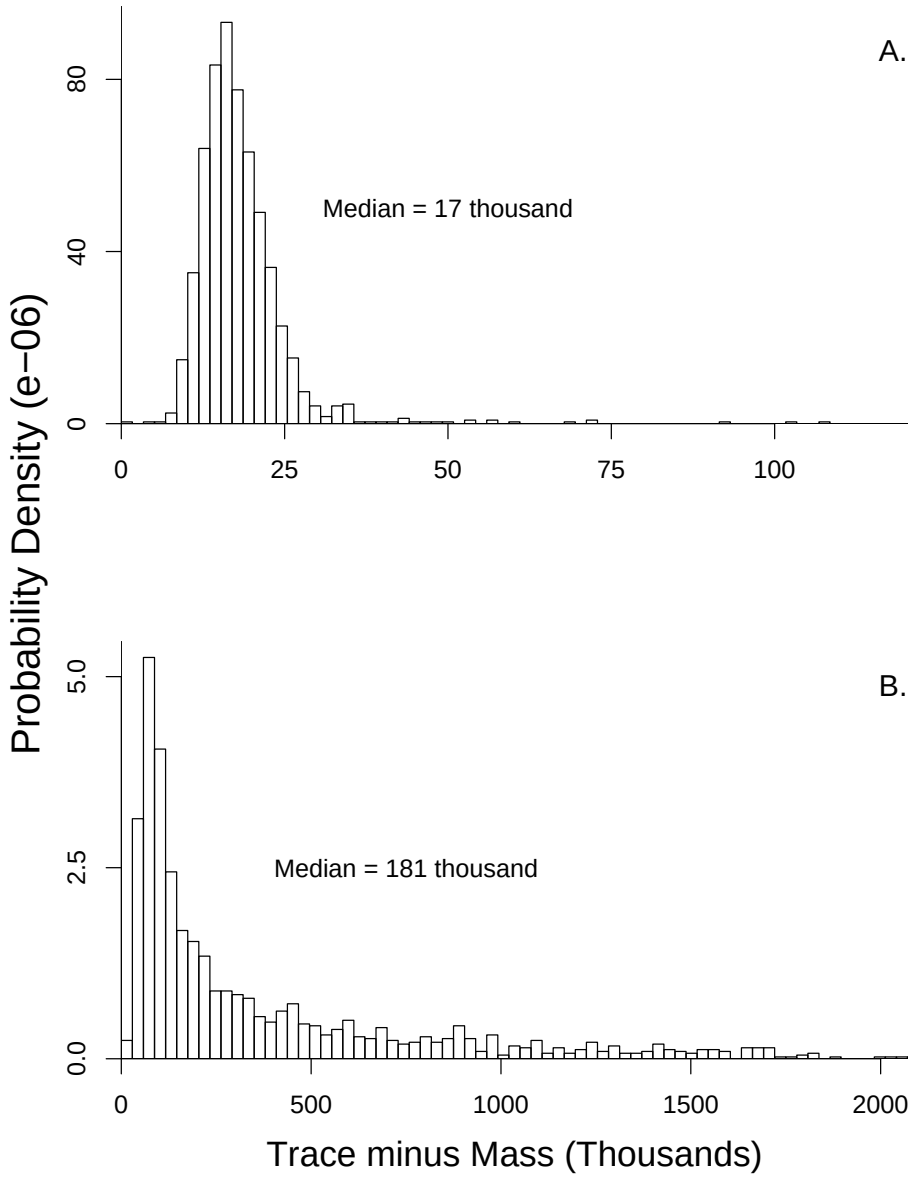
Absolute Number of Deaths Under Each Vaccination Strategy

For brevity, in the main text we show the effects of the two vaccination strategies in terms of the difference in the number of deaths between the two. In fact, the absolute number of deaths under mass vaccination is very small, especially compared to trace vaccination; moreover, most of the uncertainty in the difference in the number dead is due to uncertainty in the number dead under trace vaccination. To show this uncertainty, in Fig. 11, we show the distribution of the number of dead for both mass and trace vaccination when R_0 is estimated from Old World populations, and in Fig. 12, we show the corresponding distributions when R_0 is estimated from New World populations. As the figures show, almost all the variability in the difference in the number of dead between the two strategies is due to variability in the number of dead under trace vaccination.

Heterogeneity in Susceptibility: Impacts on $\mathrm{R_0}$ and Vaccination

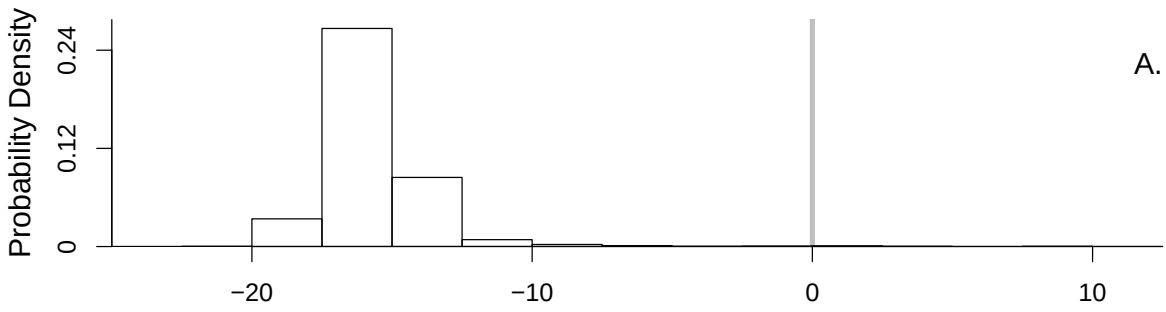
As we describe in the main text, increasing model complexity by adding an additional parameter to account for heterogeneity in susceptibility among individuals had little effect on our results. As Table 1 shows, the 95% credible interval on R_0, the equivalent of R_0 for the model with heterogeneity, has a reduced upper bound relative to the median for the model without heterogeneity, but the medians of the two distributions are nearly the same. This assumption suggests that the only meaningful change is a reduction in the tail of the distribution of R_0; indeed, Fig. 13 shows that

Fig. 9. The difference in the number of deaths between the trace and mass vaccination scenarios for an optimistic vaccination scenario. (A) is for parameters based on Old-World populations, and (B) is for parameters based on New-World populations. Note differences in scales on the axes.



the two distributions are very similar for both the previously exposed Old World and previously unexposed New World populations. As a result, there is little or no change in the distribution of the difference in the number of dead under the two vaccination strategies, as shown by a comparison of Fig. 14 to Fig. 3.

Fig. 10. The difference in the number of deaths between the trace and mass vaccination scenarios for a wildly optimistic vaccination scenario. (A) is for parameters based on Old-World populations, and (B) is for parameters based on New-World populations. The gray line represents the division between trace and mass vaccination. Note differences in scales on the axes.



This lack of difference between the outcomes of the two models is further reflected in the overall number of deaths under mass and trace vaccination for both Old World and New World populations (Fig. 15 and Fig. 16 as compared to Fig. 11 and Fig. 12, respectively). Increasing the model's complexity thus has little effect on the uncertainty in the number of dead under the two strategies.

Fig. 11. The number of deaths in Old-World populations using (A) mass vaccination and (B) trace vaccination given an outbreak in a city of 10 million. The grey histogram in each panel shows the case in which we assume that there is no uncertainty in any SEIR model parameter except R_0 ("fixed"), while the black-outlined histogram shows the case in which we instead assume that all model parameters are uncertain, and we have integrated out all the other parameters ("variable"). Note differences in scales on the axes.

Fig. 12. The number of deaths in New-World populations using (A) mass vaccination and (B) trace vaccination given an outbreak in a city of 10 million. The grey histogram in each panel shows the case in which we assume that there is no uncertainty in any SEIR model parameter except R_0 ("fixed"), while the black-outlined histogram shows the case in which we instead assume that all model parameters are uncertain, and we have integrated out all the other parameters ("variable"). Note differences in scales on the axes.

Fig. 13. Effects of prior knowledge on uncertainty in the disease transmission rate R_0 for the SEIR model and the heterogeneous SEIR model. The estimated posterior density of R_0 for Old-World (A) and New-World (B) populations are shown separately. The curved black line represents the prior distribution. The grey histogram in each panel shows the case in which we assume that there is no heterogeneity in susceptibility among individuals, while the black-outlined histogram shows the case in which we instead assume that there is heterogeneity in the population. Note differences in scales on the axes.

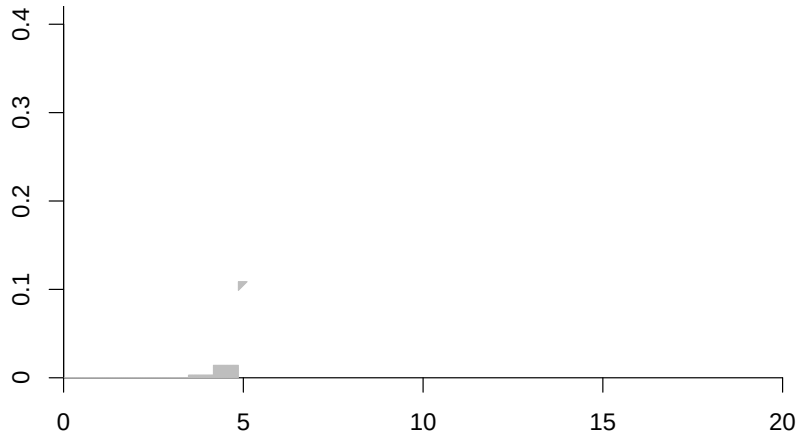


Fig. 14. Difference in the number of deaths between trace vaccination and mass vaccination strategies in a simulated population of 10 million: (A) Using parameters based on Old-World populations. (B) Using parameters based on New-World populations. Note differences in scales on the axes.

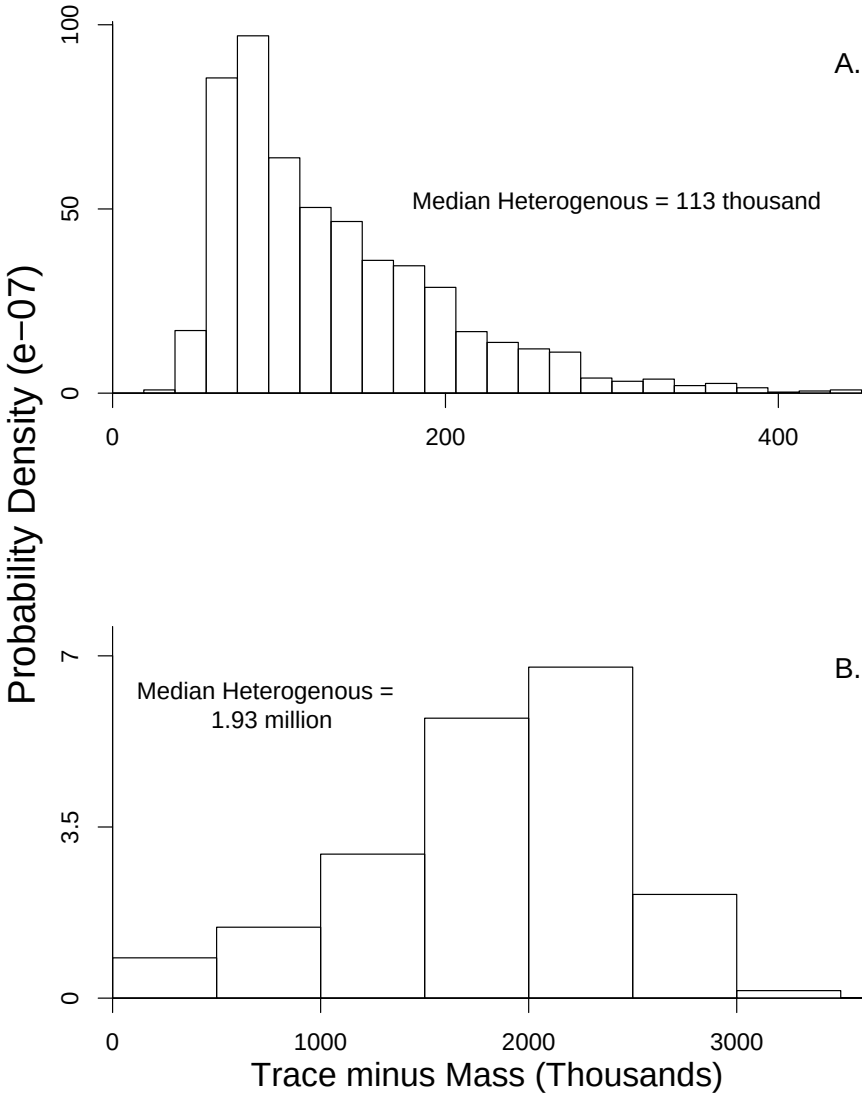


Fig. 15.

Fig. 16. The number of deaths in New-World populations using (A) mass vaccination and (B) trace vaccination given an outbreak in a city of 10 million when assuming heterogeneity in susceptibility. NriW3(in)-23384ogeneitif

A Simple Method of Including Variability in Susceptibility in the SEIR Model

In order to model the effects of heterogeneity in a population's susceptibility, we used a moment-closure approach (8–10). Under the assumption that individuals vary in their susceptibility to the disease, the SEIR model becomes:

$$\begin{aligned} \frac{@S}{@t} &= \frac{R_0}{N} SI; \\ \frac{dE}{dt} &= I\frac{}{N}\int_0^{\infty} R_0 S(R_0;t)dR_0 \quad E; \\ \frac{dI}{dt} &= E \quad I; \\ \frac{dR}{dt} &= I : \end{aligned} \tag{21}$$

The moment-closure method reduces the partial differential equation in Eq. **21** to a set of ordinary differential equations. First, we define:

$$S_j = \int_0^{\infty} R^j$$

such that each moment is a function of a higher-order moment. To approximate the higher order moments and, thus, close the moments, we assume that the coefficient of variation remains constant. This finding turns out to be a reasonable approximation for most initial distributions of susceptibility (10). If k is defined as the inverse squared coefficient of variation then:

$$k = \frac{m_1^2}{m_2 \quad m_1^2} : \tag{25}$$

Thus, by substituting Eq. **25** into Eq. **24** with respect to the first moment, we have:

$$\frac{dm_1}{dt} = \frac{I m_1^2}{k} : \tag{26}$$

Also note that Eq. **23** can be approximated with $j = 0$ as:

$$\frac{d\hat{S}}{dt} = m_1 I \hat{S} : \tag{27}$$

We then divide Eq. **26** by Eq. **27**, integrate from 0 to t and solve for m:

$$m_1 = \frac{R_0}{N} \left(\frac{\hat{S}(t)}{\hat{S}(0)} \right)^{1=k} : \tag{28}$$

remembering that at time 0, $m(0)$ is the mean of the disease reproductive rate, R_0. We now substitute Eq. **28** into Eq. **27** to derive an ordinary differential equation for the total population size75l(dan.91 Tf 419.3

where k is the inverse of the square root of the coefficient of variation of the distribution of R_0 across individual hosts. k is thus an inverse measure of the variability in susceptibility among individuals in the population. More susceptible individuals will acquire smallpox earlier in the outbreak (11). As the epidemic progresses, susceptibility of the noninfected decreases as R_0 is scaled by smaller and smaller values of $S(t)/S(0)$. The larger the value of k, the closer the model is to the homogeneous model (Eq. **1**) and models with values of $k > 200$ are nearly indistinguishable from models that assume all individuals are identical.

Note that, in numerically solving these and all other differential equations, we used an automatic second- and third-order pair step-size Runge-Kutta-Fehlberg integration method or a simple fourth- and fifth-order Runge-Kutta algorithm (Matlab; Mathworks, Natick, NA). In general, the SEIR equations (Eq. **1**) are well known for their numeric stability (i.e., the state variables change very smoothly over time) (11), and the use of both integration methods noted above perform well with such equations (12).

References

1. Bozzette, S. A., Boer, R., Bhatnagar, V., Brower, J. L., Keeler, E. B., Morton, S. C. & Stoto, M. A. (2003) *N. Engl. J. Med.* **348**, 416–425.

2. Gani, R. & Leach, S. (2001) *Nature* **414**, 748–751.

3. Efron, B. & Tibshirani, R. J. (1998) *An introduction to the boostrap* (Chapman and Hall, Boca Raton, FL).

4. Rigau-Pérez, J. G. (1982) *J. Hist. Med.* **37**, 423–438.

5. Kaplan, E. H., Craft, D. L. & Wein, L. M. (2002) *Proc. Natl. Acad. Sci. USA.* **99**, 10935–10940.

6. Eichner, M. & Dietz, K. (2003) *Am. J. Epidemiol.* **158**, 110–117.

7. Fenner, F., Henderson, D. A., Arita, I., Jezek, Z. & Ladnyi, I. D. (1998) *Smallpox and its eradication* (WHO, Geneva).

8. Dwyer, G., Elkinton, J. S. & Buonaccorsi, J. P. (1997) *Am. Nat.* **150**, 685–707.

9. Lloyd, A. (2004) *Theor. Popul. Biol.* **65**, 49–65.

10. Dushoff, J (1999) *Theor. Popul. Biol.* **56**, 325–335.

11. Anderson, R. M. & May, R. M. (1991) *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, Oxford).

12. Kreyszig, E. (1993) (Wiley).

The Sources of the Epidemic Data

Table 3. List of epidemics used in the smallpox outbreak models

Epidemic	Year	Population Size	Ref.
Boston	1721	10,565	1
Burford	1758	1,520	2
Chester	1774	12,009	1
Warrington	1773	7,000	1
Mauritius	1891	37,110	3
Saint Lawrence	1781	215	4*
Santo Domingo, New Mexico	1781	578	4*
Santa Clara	1780	627	4*
San Buenaventura	1780	501	4*
San Francisco Xavier	1780	261	4*
Rosalia	1781	133	4*
Pojoaque	1781	270	4*
Santo Domingo, Baja California	1781	119	4*

* Population sizes of initially susceptible individuals derived from the census of 1777 (5) and 1778 (6). Forecasting of population size up to the time of the epidemic was carried out using background birth and death rates from various sources (6–9).