## Supporting Text

**Age-Dependent Dose–Response Function.** An as-yet-unexplained aspect of the Sverdlovsk outbreak is the absence of any victims under the age of 24, despite the fact that children apparently were in the path of the plume (1). Fig. 8 compares the age distribution of the Sverdlovsk victims to the distribution one would expect based on Russian demography circa 1979. This anomaly has prompted several researchers to assume that the likelihood of contracting inhalation anthrax is age dependent (2–4). Nonhuman primate experiments shed no light on this anomaly because they use adult monkeys. Such age dependence may exist, but the Sverdlovsk data in Fig. 8 does not show a pronounced age dependence relative to what one would expect based on Russian demography because the large number of cases in the 40- to 50-year-old age range can be explained by the age of the workers at the pipe factory, which constituted the largest cluster of victims (18) in the Sverdlovsk data. Moreover, in general, children are more susceptible to infectious diseases because of their immature immune system, not less susceptible, as suggested by the Sverdlovsk data. Breathing rates, lung morphology, and the exact demography of the Chkalovsky district of Sverdlovsk might explain some difference in the infection rate between children and adults, but one would not expect the complete absence of children in a sample of 70 victims.

Fig. 9, on the other hand, does show a clear age dependence for the duration of the incubation period among male victims, but no pronounced age dependence for the incubation period for female victims. The 95% confidence intervals are shown for a linear regression of the incubation period against the victim's age.

**Reproducing the Spatial Distribution of Sverdlovsk Victims.** Fig. 10 illustrates pictorially the plumes for all four anthrax models examined in this paper, where the release amount in each case is that required to produce 81 infected individuals (see Fig. 2). The cumulative distribution of victims as a function of the exposed dose, derived by simulating the release of an amount of anthrax sufficient to produce 81 infected victims for each model, is shown in Fig. 11.

**Modeling the Incubation Period Distribution.** The following discussion gives a brief introduction to the way in which the incubation distribution is modeled in this analysis. More details can be found in the references. At the very least, this discussion allows the reader to understand the parameters listed in Table 1.

After the pioneering work of Sartwell (5), it is common to assume that incubation periods follow a log-normal distribution in time, with the cumulative incubation distribution, $F(t)$, given by

$$F(t) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)\int_0^t \left(\frac{1}{x}\right)\exp\left(-\frac{(\ln(x)-\ln(M))^2}{2\sigma^2}\right)dx,$$

where $M$ is the median incubation period and $\sigma$ is the SD. The dose dependence of the incubation period in model A1 is given by

$$M = \alpha + \beta\log(D),$$

where $D$ is the exposed dose and $\alpha$ and $\beta$ are parameters associated with the log-linear model. The SD is given by a similar equation,

$$\sigma = \gamma + \delta\log(D),$$

where $\gamma$ and $\delta$ are parameters. The parameter values for the median and SD in model A1 are obtained from a least-square fit to the Sverdlovsk data at the low-dose end of the spectrum and nonhuman primate data at the high-dose end (6–9), as shown in Figs. 12 and 13. These fits use only the solid data points in the figures. The open data points are useful for comparison, but these data either contain too few nonhuman primates or do not include the raw data upon which to construct the complete time-to-death distribution and, hence, are not used in the fits.

Note that the nonhuman primate data are given for the time from exposure to death. Consequently, one must estimate the equivalent human incubation period from the nonhuman primate time to death data. Note also that the data points in Fig. 12 do not match exactly the points in Fig. 4 because the median values in Fig. 12 are derived by finding the maximum likelihood fit for a log-normal distribution to the Sverdlovsk and nonhuman primate data, whereas the median values in Fig. 4 are derived by fitting a modified Brookmeyer model to the Sverdlovsk and nonhuman primate data. These differences are not significant.

Clearly, one would have greater confidence in these fits if more data were available in the mid-dose range and if these data were not mixed between human and nonhuman primate exposures. The 95% confidence intervals in both figures imply that there is considerable uncertainty in the parameter values derived from these fits. This uncertainty has important policy implications because it translates into an uncertainty in the early-time tail of the incubation period distribution and, hence, in the likelihood that victims will appear symptomatic soon after exposure. This uncertainty, in turn, affects one's estimate of when astute physicians might diagnose the first patient with inhalation anthrax and, hence, provide warning that a bioterror attack has occurred absent any warning signal from environmental sensors, as discussed in the main article.

The parameter values for models B and C can be found in the references (4, 10, 11). Model D is a modified version of a model developed by Brookmeyer *et al.* (unpublished data; ref. 12), wherein the incubation period is described by two competing risks, an exponentially distributed probability of spore germination and an exponentially distributed probability of spore clearance from the lungs. The resulting probability that a spore germinates before it is cleared from the lungs is given by

$$F(t) = \frac{1-(1-p)^{(1-e^{-(\lambda+\theta)t})}}{p}, \quad \textbf{[1]}$$

where $p$ is the probability of becoming infected as a function of the number of spores to which a victim is exposed and is given by,

$$p = 1 - \exp\left(\frac{-D\lambda}{\theta + \lambda}\right), \quad \textbf{[2]}$$

where $D$ is the exposed dose in spores, $\lambda$ is the rate at which spores germinate, and $\theta$ is the rate at which spores are effectively removed from the lungs. If a time delay is included in the model to account for bacterial growth before the onset of symptoms, then the incubation period distribution can be written as the convolution of the probability that a spore germinates before being rendered nonviable and the probability that the time delay has a given duration, as discussed in greater detail (unpublished data). The resulting equation is

$$F(t) = \int_0^t f(t-s)g(s)ds, \quad \textbf{[3]}$$

where $f(t - s)$ is the probability that a spore germinates before being destroyed, given in Eq. **1**, and $g(s)$ is the probability that the time delay has duration $s$. The delay time can be modeled as a "lag phase" followed by an "exponential growth phase." Thus, one can represent the time from spore germination to the onset of symptoms, $T$, as follows,

$$T = t_{lag} + \left(\frac{t_2}{\ln(2)}\right) \cdot \ln\left(\frac{N_{thresh}}{D}\right), \quad \textbf{[4]}$$

where $t_{lag}$ is the lag time (assumed to be 1 h in this analysis), $t_2$ is the bacterial doubling time, $N_{thresh}$ is the threshold number of bacillus at which symptoms appear, and $D$ is the initial number of *Bacillus*, which equals one regardless of the exposed dose because the first spore to germinate starts the bacterial growth phase. In modeling the delay time as a stochastic variable, one should chose a distribution that accurately represents bacterial

growth. Because bacterial growth is governed by multiple factors in the host, it is reasonable to assume that $g(s)$ follows a log-normal distribution, i.e.,

$$g(s) = \left(\frac{1}{\sigma_g \sqrt{2\pi}}\right)\left(\frac{1}{s}\right)\exp\left(-\frac{(\ln(s) - \ln(M_g))^2}{2\sigma_g{}^2}\right), \quad \textbf{[5]}$$

where $M_g$ is the median value of the time delay due to bacterial growth and $\sigma_g$ is the SD. This distribution has the virtue that its value is 0 when $s$ equals 0, it peaks near the median value, $M_g$, and then tapers off for $s > M_g$. The incubation distribution is then given by Eq. **1**–**3** and **5**, with $M_g$ given by Eq. **4**. Unfortunately, the integral in Eq. **3** must be solved numerically, which adds a computational burden to the use of this model. The parameter values for model D are found from a maximum likelihood fit to the Sverdlovsk incubation data, assuming the model D dose–response curve and, consequently, that the Sverdlovsk victims were exposed to a mean (geometric) dose of 360 spores, as illustrated in Fig. 4.

The median incubation period for model A2 has the same functional form as that given in model D, with the exception that model A2 assumes the model A dose–response curve and, consequently, that the Sverdlovsk victims were exposed to a mean (geometric) dose of 2.4 spores, as illustrated in Fig. 4. Consequently, the parameter values for model A2 differ somewhat from those for model D. The SD for model A2 is the same as that in model A1 and, hence, model A2 is a hybrid of models A and D.

Models A2 and D simplify what is obviously a complex host response to bacterial infection because both models assume a simple bacterial threshold for symptom onset equal to $10^9$ *Bacillus anthracis*. The selected threshold value affects the estimate of the bacterial doubling time obtained when fitting the model to the Sverdlovsk incubation data. Varying the threshold value parametrically produces a range of estimates for the bacteria doubling time, as shown in Fig. 14.

Using these incubation period models and the distribution of people infected versus dose shown in Fig. 11, one can calculate the expected incubation period distribution for the Sverdlovsk outbreak by taking into account the fact that the victims were not all exposed to the same dose. The result is shown in Fig. 15. Models A1, A2, C, and D all have distributions that reproduce the Sverdlovsk data, which is not surprising because the parameters for these models were derived from fits to the Sverdlovsk data. The model A1 and A2 fits are nearly identical, even though model A2 provides a better fit to the high-dose nonhuman primate incubation period data, as discussed in the main article, because the high-dose incubation period distribution is not relevant for the range of doses to which the Sverdlovsk victims were exposed. Model C has a longer tail because it is based on a constant median incubation period of 11 days, independent of dose. Model D also gives a very good fit to the Sverdlovsk data. The average residuals for models A1, A2, C, and D are 1.032, 1.028, 1.12 and 1.055, respectively; indicating that models A1 and A2 are virtually indistinguishable and that models A1 and A2 provide marginally better fits to the Sverdlovsk data than model D. Clearly, accurate data for the median incubation period at the $ID_{50}$ dose would help distinguish between models A1 and A2. Model B does not fit the Sverdlovsk data.

1. Guillemin, J. (1999) *Anthrax: Investigation of a Deadly Outbreak* (Univ. of California Press, Berkeley).

2. Meselson, M., Guillemin, J., Hugh-Jones, M., Langmuir, A., Popova, I., Shelokov, A. & Yampolskaya, O. (1994) *Science.* **266,** 1202–1208.

3. Webb, G. F. & Blaser, M. J. (2002) *Proc. Nat. Acad. Sci. USA* **99,** 7027–7032.

4. Wein, L. M., Craft, D. L. & Kaplan, E. H. (2003) *Proc. Nat. Acad. Sci. USA* **100,** 4346–4351.

5. Sartwell, P. (1950) *Am. J. Hyg.* **51,** 310–318.

6. Henderson, D. W., Peacock, S. & Belton, F.C. (1956) *J. Hyg.* **54,** 28–36.

7. Friedlander A. M., Welkos, S. L., Pitt, M. L., Ezzell, J. W., Worsham, P. L., Rose, K. J., Ivins, B. E., Lowe, J. R., Howe, G. B., Mikesell, P., *et al.*(1993) *J. Infect. Dis.* **167,** 1239–1243.

8. Ivins, B. E., Fellows, P. F., Pitt, M. L.M., Estep, J. E., Welkos, S. L., Worsham, P. L., & Freidlander, A. M. (1996) *Salisbury Medical Bulletin-Special Supplement* **87,** 125–126.

9. Gleiser, C.A., Berdjis, C.C., Hartman, H.A., & Gochenour, W.S. (1963) *Brit. J. Exper. Pathol.*, **44,** 416–426.

10. Anno, G. H. & Bloom, R. (1998) *Anthrax Effects Model* (Pacific Sierra Research Corp., Santa Monica, CA) PSR Rep. 2773B.

11. Anno, G. H. (2001) *Susceptibility of Mammalian Species to Inhalation of Aerosolized Bacillus Anthracis Spores* (Veridian Systems Division, Santa Monica, CA).

12. Brookmeyer, R., Johnson, E. & Barry, S. (2005) *Stat. Med.* **54,** 531–542.