*Risk Analysis, Vol. 22, No. 6, 2002*

# A Risk Analysis for Airborne Pathogens with Low Infectious Doses: Application to Respirator Selection Against *Coccidioides immitis* Spores

**Mark Nicas[1]* and Alan Hubbard[1]**

Probability models incorporating a deterministic versus stochastic infectious dose are described for estimating infection risk due to airborne pathogens that infect at low doses. Such pathogens can be occupational hazards or candidate agents for bioterrorism. Inputs include parameters for the infectious dose model, distribution parameters for ambient pathogen concentrations, the breathing rate, the duration of an exposure period, the anticipated number of exposure periods, and, if a respirator device is used, distribution parameters for respirator penetration values. Application of the models is illustrated with a hypothetical scenario involving exposure to *Coccidioides immitis*, a fungus present in soil in areas of the southwestern United States. Inhaling *C. immitis* spores causes a respiratory tract infection and is a recognized occupational hazard in jobs involving soil dust exposure in endemic areas. An uncertainty analysis is applied to risk estimation in the context of selecting respiratory protection with a desired degree of efficacy.

**KEY WORDS:** Airborne pathogens; infection risk; coccidioidomycosis

## 1. INTRODUCTION

It is estimated that five million U.S. workers are exposed to air contaminants at levels requiring the use of respiratory protective devices at least part of the time.[1] Respirators are most frequently worn against chemical toxicants in the form of vapors, gases, and nonviable aerosols. The general decision strategy for selecting respirators against chemicals is well established,[2–4] and although it has numerous steps, the overall idea is simple—one chooses a respirator predicted to keep the inspired level at or below the chemical's occupational exposure limit. Airborne pathogens are ignored in published decision strategies,[2–4] but the importance of selecting appropriate respirators is implicit in efforts by the Centers for Disease Control and Prevention to promote domestic preparedness against bioterrorism.[5]

Selecting an air-purifying respirator against an airborne pathogen is more complex than against a chemical agent because there are no exposure limits for the number of organisms that may be inhaled or for infection risk. Past recommendations from public health agencies for respirator use against Hantavirus[6] and *Mycobacterium tuberculosis*[7] have been based on expert opinion and not on quantitative risk analysis. More recently, a strategy was described that applied to a broad array of infectious aerosols, but it was also a qualitative approach involving unspecified criteria for acceptable risk.[8]

For pathogens exhibiting high infectious dose values, we have described a probability framework for estimating infection risk, and for comparing the efficacy of different respirators in reducing risk.[9] The example used was *Bacillus anthracis*, which causes

1 Center for Occupational and Environmental Health, School of Public Health, University of California, Berkeley, CA.
*Address correspondence to Mark Nicas, Ph.D., CIH, School of Public Health, Room 140 Warren Hall, University of California, Berkeley, CA 94720; mnicas@uclink4.berkeley.edu.


0272-4332/02/1200-1153$22.00/1 

inhalation anthrax and is the leading candidate agent for domestic bioterrorism. The median infectious inhalation dose is 8,000–10,000 spores as estimated by the U.S. Defense Intelligence Agency,[10] although published data for primates and guinea pigs indicate the median infectious dose is closer to 60,000 spores.[11] Our purpose here is to extend the framework to airborne pathogens that cause infections at low inhalation doses and to illustrate how uncertainty analysis may be applied to estimating infection risk. The example used is *Coccidioides immitis*, a soil fungus that causes a human infection termed coccidioidomycosis.[12] Infectious doses of *C. immitis* are in the range of 1–10 spores.

To our knowledge, the only previous quantitative risk model for airborne pathogens was a one-hit model proposed by William Wells and Richard Riley in the context of *M. tuberculosis* and influenza infection.[13,14] These investigators used a deterministic "quantum" unit to compute infection risk, where a quantum was defined as that number of organisms required to initiate infection. We will show that when a deterministic infectious dose is more than one organism, the Wells-Riley model is not appropriate.

## 2. INDIVIDUAL RISK

Because an individual may have multiple exposure periods, it is necessary to consider how infection risk accumulates. Computing cumulative risk requires specifying: (1) an appropriate probability model for the infectious inhalation dose; (2) the ambient air concentrations of the pathogen, $C_A$ (#/m$^3$); (3) the respirator penetration values, P (0–1); (4) the breathing rate $Q_V$ (m$^3$/hr); (5) the duration of an exposure period, T (hr); and (6) the anticipated number of exposure periods, M. Because such complete information will almost never be available, an uncertainty analysis of the model prediction should be performed. For this discussion, a number of default assumptions are employed.

### 2.1. The Infectious Dose

Discussing the infectious dose involves considering the alternative frameworks of determinism versus stochasticism. In a deterministic construct, it is assumed that a specified number of organisms B must deposit at the appropriate site in the lungs in an appropriate time interval to initiate infection. For example, if an individual's infectious dose is $b = 10$ organisms, infection occurs only if 10 or more organisms are received. Interindividual variability in susceptibility pertains to the distribution of infectious dose values across a cohort. If all hosts receive D organisms, the expected risk of infection (i.e., the infection risk for a randomly selected host) is the fraction of the cohort for whom $b \leq$ D.

The stochastic construct assumes that only one organism is necessary to initiate infection, with each organism having an independent probability of success in that endeavor. It is assumed that the probability of success does not vary across the organisms received by a given individual, although it may vary across hosts.[15] If r denotes the probability of success per organism received by a host, and if D organisms are received, infection occurs with probability: $1 - (1 - \mathrm{r})^{\mathrm{D}}$. If r is constant across all hosts who receive D organisms, each host has the same infection risk. However, if the probability of success per organism is a random variable $\mathrm{r_i}$ across hosts, and if all hosts receive D organisms, the expected risk of infection is $1 - \mathrm{E}[(1 - \mathrm{r_i})^{\mathrm{D}}]$. A traditional probability model for the $\mathrm{r_i}$ is a standard beta distribution with parameters $\alpha$ and $\beta$.[15–17] In this setting, if all hosts receive D organisms, the expected risk of infection is: $1 - \{\Gamma(\alpha + \beta)\Gamma(\beta + \mathrm{D})/[\Gamma(\beta)\Gamma(\alpha + \beta + \mathrm{D})]\}$.

We are not aware of published studies for airborne pathogens that have compared the deterministic and stochastic models. In our earlier probability framework for inhalation anthrax,[9] we invoked determinism for two reasons. First, given the relatively large numbers of spores required to infect, it seemed more plausible that a host's defense mechanisms had to be "overwhelmed," in the sense of exceeding a threshold dose, as opposed to the circumstance that each spore had an independent probability of infection on the order of $10^{-5}$. Second, our analysis of data on guinea pig mortality due to *B. anthracis* spore inhalation[11] showed that a deterministic log-normal distribution model provided a better fit than did several stochastic models. Although a stochastic model provided a slightly better fit for monkey mortality data, in each exposure group the number of monkeys (eight) was one-fifth the number of guinea pigs (40). These analyses are summarized in Appendix 1. Overall, the anthrax data are too sparse to permit a distinction between determinism and stochasticism. For the present analysis involving the infectivity of small numbers of organisms, as pertains to *C. immitis*, we consider both types of infectious dose models.

For simplicity, we assume that all inhaled organisms deposit in the target region of the lungs, such that we need not distinguish between the

inhaled dose and the deposited dose. We recognize that these two doses are *not* equivalent, and that only pathogens that deposit can infect the host. The reason for the assumption is that published laboratory studies on inhalation exposure report only the presumed inhalation dose and not the deposited dose. Further, the deposition probability can be factored into the analysis, where that probability depends on the aerodynamic diameters of the particles carrying the pathogen; particle aerodynamic diameters can vary with the pathogen and the associated environmental media.

An issue applying to the deterministic but not the stochastic infectious dose construct is cumulative dose. That is, the number of inhaled (strictly, deposited) pathogens in the lungs might accumulate across exposure periods close in time and build to a level exceeding the individual's infectious dose. Because available evidence suggests that for some pathogens the number of organisms in the lungs could increase across exposure periods occurring within several days, we assume that a cumulative dose effect is possible.

### 2.2. Other Variables

Several simplifying assumptions are used. We assume that an airborne particle can carry only one pathogen. This idea is reasonable if the initial target tissue is the pulmonary region and if the pathogen is in the 1–5 $\mu$m size range, as is true for *C. immitis* spores. The ambient exposure level $C_A$ is defined as a constant airborne concentration of pathogens over a T-hr interval. For an individual, we assume that $C_A$ is log-normally distributed across exposure periods. Given the overall lack of sampling data for infectious aerosols, there is no strong justification for assuming any specific distribution for $C_A$. However, occupational aerosol exposure levels are usually well described by log-normal distributions,[18] so it is reasonable to invoke that distribution for airborne pathogens.

Similarly, the fractional respirator penetration P is defined as a constant value over a T-hr interval, and is reasonably assumed to be log-normally distributed and independent of the $C_A$ values.[19,20] Penetration is due primarily to inward leakage of air around the respirator's sealing perimeter. In a strict sense, P cannot be log-normally distributed because it cannot exceed one. However, for the types of respirators considered here, there is negligible probability associated with P greater than one.

We assume that an individual's breathing rate $Q_V$ is constant, and for the *C. immitis* example, $Q_V = 1.75\ m^3/hr$. The latter value corresponds to a medium work rate,[2] and is higher than the typically assumed $1.0\ m^3/hr$, which corresponds to a light work rate. Activities performed by individuals with the potential for *C. immitis* exposure (agricultural workers, construction workers) would be more rigorous than "light" work. We also assume that each exposure period is the same duration; T = 2 hr is a reasonable value.

### 2.3. Cumulative Risk with No Cumulative Dose Effect

The lack of a cumulative dose effect in the context of determinism means that the infectious dose must be received in one exposure period. For example, if a person's infectious inhalation dose is $b = 10$, receiving nine organisms in each of two or more consecutive exposure periods does not cause infection. The lack of a cumulative dose effect is unrealistic where two or more periods are close in time and is likely invalid for *C. immitis* spore exposures occurring within a day or two. Rather, the scenario is offered primarily as a contrast to the situation in which the dose can accumulate. Further, the scenario would pertain to exposure periods far removed in time, such that no viable organisms inhaled during previous exposures are likely to remain.

Let $q_i$ denote the probability of not being infected during the $i^{th}$ exposure period, and let M denote the number of periods. The cumulative infection risk over M periods conditioned on a particular set of $q_i$ values, denoted $R^*_M$, is:

$$R^*_M = 1 - \prod_{i=1}^{M} q_i \tag{1}$$

If the $q_i$ are independent and identically distributed random variables across the M periods, the unconditional cumulative infection risk is given by:

$$R_M = 1 - (E[q])^M \tag{2}$$

where E[q] is the expected value of the random variable $q_i$. Equation (2) follows from the property that the expected value of the product of two or more independent random variables is the product of their respective expectations.

E[q] is related to $C_A$, P, $Q_V$, and T in the following manner. The inhalation dose during the $i^{th}$ exposure period, $D_i$ (# organisms), can be written as:

$$D_i = C_{A,i} \cdot P_i \cdot Q_V \cdot T \tag{3}$$

where $C_{A,i}$ and $P_i$ denote the ambient concentration and respirator penetration value, respectively, during the $i^{th}$ exposure period. $P_i = 1$ if a respirator is not worn. The description of the actual number of organisms inhaled depends on the magnitude of $D_i$. If the dose is usually a small number, say, 10 or less, $D_i$ should be treated as a discrete random variable even though $C_{A,i}$ and $P_i$ are treated as continuous random variables. In this case, the integer dose for the $i^{th}$ exposure period is modeled as a Poisson random variable with expected value $\lambda_i$ equal to the RHS of Equation (3), or:

$$\Pr[D_i = d] = \frac{(\lambda_i)^d \exp(-\lambda_i)}{d!} \tag{4}$$

If $B = b$ denotes the individual's deterministic infectious dose value, the probability of not being infected during the $i^{th}$ exposure period is:

$$q_i \mid B = b = \Pr[D_i < b] = \sum_{d=0}^{b-1} \frac{(\lambda_i)^d \exp(-\lambda_i)}{d!} \tag{5}$$

Unlike the actual dose received, the expected dose $\lambda_i$ (the parameter of the Poisson probability mass function) is a continuous variable. The "unconditional" probability $E[q \mid B = b]$ is found by the double integration of $q_i \mid B = b$ across the respective ranges of $C_A$ and P, which constitutes taking the expectation by conditioning on the value of $\lambda_i$:

$$E[q \mid B = b] = \sum_{d=0}^{b-1} p_d \tag{6}$$

where

$$p_d = \int_{C_A} \int_{P} \frac{(\lambda)^d \exp(-\lambda)}{d!} f_{C_A}(c_A) f_P(p)\, dc\, dp$$

where $\lambda = C_A \cdot P \cdot Q_V \cdot T$, and $f_{C_A}(c_A)$ and $f_P(p)$ denote the probability density functions for, respectively, $C_A$ and P. Without respirator use, integration over P would not occur.

For a stochastic infectious dose with probability of success per organism r in a given host, $E[q \mid r]$ is computed by:

$$E[q \mid r] = \sum_{d=0}^{\infty} (1 - r)^d\, p_d \tag{7}$$

where $p_d$ is defined in Equation (6). Note that Equation (7) yields a quantity slightly different than the term $e^{-\lambda \cdot r}$ in the traditional exponential model.[15] In the latter, $\lambda$ is a fixed parameter for a single exposure, but here $\lambda$ is a random variable across exposure periods.

### 2.4. Cumulative Risk with a Cumulative Dose Effect

Consider that the inhalation dose (strictly, the deposited dose) can accumulate across exposures close in time. Let $S_M$ denote the sum of the doses over M consecutive exposure periods close in time, or:

$$S_M = \sum_{i=1}^{M} D_i \tag{8}$$

where $D_i$ is the inhaled dose for the $i^{th}$ exposure period. Given a deterministic infectious dose $b$ and no loss of viable organisms from previous periods, the individual's cumulative infection risk is:

$$R_M = 1 - \Pr[S_M < b] \tag{9}$$

As before, if the inhaled dose is a small number, $D_i$ is modeled as a Poisson random variable with expected value $\lambda_i$ equal to the RHS of Equation (3). Again, although the $D_i$ are all Poisson variables, they are not identically distributed due to variability in the $\lambda_i$. If the infectious dose and the number of exposure periods are small, it is practical to determine $R_M$ as follows. In a random exposure period, the probability of inhaling $d$ pathogens is defined by the term $p_d$ in Equation (6). Consider an individual for whom $b = 3$. For $m = 1$, infection risk is: $1 - (p_0 + p_1 + p_2)$. For $m = 2$, infection risk is: $1 - (p_0^2 + 2 \cdot p_0 \cdot p_1 + 2 \cdot p_0 \cdot p_2)$, where the sum in the parentheses accounts for the number of ways leading to $S_2 < 3$. For $m = 3$, infection risk is: $1 - (p_0^3 + 3 \cdot p_0^2 \cdot p_1 + 3 \cdot p_0 \cdot p_1^2 + 3 \cdot p_0^2 \cdot p_2)$, where the sum in the parentheses accounts for the number of ways leading to $S_3 < 3$; and so forth. Because this procedure is computationally tedious for large $b$ or $m$, Monte Carlo simulation is more readily used to determine the probability distribution of $S_M$.

For simulation, the geometric means and geometric standard deviations of the log-normal $C_A$ and log-normal P distribution are specified, along with values for $Q_V$ and T. For a given simulation, a value is randomly drawn from the $C_A$ distribution, and a value is randomly drawn from the P distribution. The corresponding $\lambda$ value is computed by Equation (3), and a Poisson random variable D with expected value $\lambda$ is generated. This procedure is repeated M times, and the sum of these random D values equals $S_M$. Fifty-thousand $S_M$ values are so generated to obtain a stable distribution. By inspection (via computer), the fraction of $S_M$ values less than $b$ equals $\Pr[S_M < b]$.

Even if the $D_i$ variables were identically distributed, a decrease in the number of previously deposited organisms in the lungs would be expected due

to respiratory tract clearance mechanisms. Therefore, the number of viable organisms in the lungs after M exposure periods should be less than the sum of the $D_i$ values. Although one could posit a function describing the decrease in viable organisms over time, for simplicity, we do not consider respiratory tract clearance. Further, experiments with radio-labeled *B. subtilis* spores showed that nearly 100% of spores that deposited in the alveolar region of guinea pigs were present 24 hours later.[21] Because our example for similarly sized *C. immitis* spores involve exposures occurring over two days, the risk estimates should be little affected by ignoring respiratory tract clearance.

## 3. COHORT INFECTION RISK

In essence, for each cohort member one specifies: (1) the distributions of $C_A$ and P, (2) the deterministic infectious dose B, or the stochastic success probability per organism r, and (3) the values of $Q_V$ and T. Next, the individual's cumulative infection risk $R_M$ is computed by one of the methods previously outlined. The mean of the individual $R_M$ values equals $CR_M$. If $Q_V$ and T are constant across individuals and time, the cohort analysis must still account for interindividual variability in B or r. There may also be interindividual variability in mean $C_A$ values and in mean respirator P values, but only variability in B and r are considered here.

### 3.1. Cumulative Risk with No Cumulative Dose Effect

Assume that everyone experiences the same respective log-normal distributions of $C_A$ and P with M exposure periods. For small deterministic infectious dose values, a probability mass function $\Pr[B = b]$ models the discrete infectious dose distribution across hosts. Let $E[q \mid B = b]$ be the probability of no infection during a random exposure period given $B = b$. The cohort infection risk is:

$$\mathrm{CohR_M} = \sum_{b=1}^{\infty} (1 - (E[q \mid B = b])^M) \cdot \Pr[B = b] \tag{10}$$

For a stochastic infectious dose with r constant across all hosts, $\mathrm{CohR_M} = 1 - (E[q \mid r])^M$, where $E[q \mid r]$ is specified by Equation (7). Because everyone is equally susceptible, the cohort risk is any member's risk. In the alternative, if r is a random variable across hosts and is described by a standard beta distribution with parameters $\alpha$ and $\beta$, the cohort infection risk is:

$$\mathrm{CohR_M} = \int_0^1 (1 - (E[q \mid r])^M) \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \times r^{\alpha - 1} (1 - r)^{\beta - 1} \, dr \tag{11}$$

where $E[q \mid r]$ is again specified by Equation (7).

### 3.2. Cumulative Risk with a Cumulative Dose Effect

Analogous procedures are used where a cumulative dose effect pertains. Again assume that everyone experiences the same respective log-normal distributions of $C_A$ and P, that everyone has M exposure periods, and that B has a discrete probability mass function. Let $\Pr[S_M \geq B \mid B = b]$ denote the probability that the cumulative inhaled dose exceeds the infectious dose value over M periods given $B = b$. The cohort infection risk is:

$$\mathrm{CohR_M} = \sum_{b=1}^{\infty} \Pr[S_M \geq B \mid B = b] \cdot \Pr[B = b] \tag{12}$$

## 4. COCCIDIOIDES IMMITIS

### 4.1. Background

Coccidioidomycosis is an infection, typically limited to the respiratory tract, caused by inhaling *C. immitis* spores. The fungus infects soil in areas of the southwestern United States, primarily California, Arizona, New Mexico, and Texas.[12,22,23] Aerosolization of spores occurs via wind currents and human activities (plowing, digging, excavation) that disturb soil. Overall, 60% of infected persons are asymptomatic or have mild symptoms of an upper respiratory tract infection that resolve spontaneously. In the remaining 40%, a lower respiratory tract infection develops with more severe symptoms, such as fever, cough, weakness, and difficulty in breathing; acute illness usually resolves without medical therapy but may last for several weeks. In less than 1% of persons, the organism disseminates outside the lungs and leads to meningitis and infections of bones, joints, skin, and soft tissues. The risk of disseminated disease varies among ethnic groups, with the risk being 10-fold greater among African Americans and Filipinos than among Caucasians.[22]

In 1998, there were 719 cases reported in California,[24] although underreporting of cases may

be substantial. Over a five-year period in the early 1980s, the average incidence was estimated to be 30% greater than the number reported to the California Department of Health Services (CDHS); in one county endemic for *C. immitis*, the average incidence was 600% greater than reported to the CDHS.[12] Disease risk is increased for occupations with exposure to soil dust. Among all recognized occupational cases reported to the California Department of Industrial Relations in the period 1960–1972, 17% involved agriculture, 21% involved construction laborers, and 34% involved state and local government.[12] A variety of governmental jobs can involve soil dust exposure (e.g., surveyors, construction engineers, zoologists).

### 4.2. The Infectious Dose

Published data on airborne infection by *C. immitis* spores are sparse. Approximately 40 years ago, U.S. Army researchers conducted experiments in which small groups of Rhesus monkeys (*Macaca mulatta*) were exposed to different spore aerosol concentrations.[25–27] These data are summarized in Table I. With one exception, the listed dose is the estimated number of spores inhaled by a monkey. Inhalation of 10 spores infected all animals, whereas the severity of the infection sequela was dose dependent. For example, among five monkeys inhaling 50 spores, the lung lesions were localized with little destruction of lung tissue and dissemination to lymph nodes, and no deaths occurred; in three of five monkeys inhaling 300 spores, there was extensive pulmonary disease resulting in death.[27]

**Table I.**

| Dose | Number Exposed | Infections | Mortality | Reference |
|---|---|---|---|---|
| 10 | 5 | 100%[a] | NR | 25 |
| 10[b] | 10 | 100% | 30% | 26 |
| 50 | 5 | 100% | 0 | 27 |
| 100 | 5 | 100%[a] | NR | 25 |
| 300 | 5 | 100% | 60% | 27 |
| 1,000 | 5 | 100%[a] | NR | 25 |
| 10,000 | 5 | 100% | 80% | 27 |

[a] It is inferred from the description in Reference 25 that all exposed monkeys were infected, although the actual data were not presented.
[b] Intratracheal inoculation of spores.
NR = not reported.

These data do not permit distinguishing between a deterministic versus stochastic infectious dose model. A plausible deterministic model for humans is the following. Human susceptibility is thought to lie between that of the monkey and the dog, and studies have shown that intratracheal inoculation of 10 *C. immitis* spores infects 100% of dogs.[26] The inhalation dose infecting nearly 100% of exposed humans should also be ≤10 spores, and it is thought that as few as one or two spores can initiate infection in many individuals.[12] Therefore, posit that the infectious dose distribution is geometric with parameter $\theta$. The proportion of individuals requiring an integer dose $b$ for infection is given by: $\Pr[B = b] = \theta(1-\theta)^{b-1}$.

One way to estimate bounds for $\theta$ is to find a set of $\theta$ values such that that the probability is at least 95% that 5/5 randomly selected persons are infected given an inhalation dose of 10 spores. Again, 5/5 monkeys each inhaling 10 spores were infected (see Table I). According to the geometric distribution model, the population fraction for which the infectious dose is 10 spores or less, $\Pr[B \le 10]$, is:

$$\Pr[B \le 10] = \sum_{b=1}^{10} \theta(1-\theta)^{b-1} \tag{13}$$

Among five randomly selected persons, the probability that all five have an infectious dose value ≤10 spores is the binomial probability: $(\Pr[B \le 10])^5$. Therefore, find $\theta$ such that $(\Pr[B \le 10])^5 \ge 0.95$. This criterion is satisfied for $\theta \ge 0.37$. The upper bound $\theta = 1$ signifies that all individuals are infected by one spore.

In the stochastic framework, with r constant across all hosts, the analogous procedure is to find a set of r values such that: $[1-(1-r)^{10}]^5 \ge 0.95$. This criterion is satisfied by $r \ge 0.37$. Interestingly, if $r_i$ is a random variable across hosts, a standard beta distribution on $[0, 1]$ for the $r_i$ is inconsistent with the above criterion, because no $r_i$ can be less than 0.26. To explain, the lowest r value permitted would correspond to 4/5 hosts each having $r = 1$. To satisfy the criterion that 5/5 hosts who each receive 10 spores are infected with at least 95% probability, the r value for the fifth host cannot be less than 0.26.

## 5. APPLYING THE PROBABILITY MODELS

### 5.1. A Hypothetical Example

Consider a group of road construction workers who disturb soil in an area endemic for *C. immitis*.

There are no published data for airborne spore levels during road construction, but based on calculations not shown here we estimate the concentration can be on the order of 1–10 per $m^3$. Assume the airborne spore concentration is log-normal with $GM[C_A] = 0.54$ per $m^3$ and $GSD[C_A] = 3$, such that $E[C_A] = 1$ per $m^3$. Assume that each worker's breathing rate is $Q_V = 1.75$ $m^3$/hr, and that each worker has available a Type N100 half-mask respirator to wear during eight two-hour periods over two days while exposed to soil dust. Assume that each worker's respirator penetration values are log-normal with $GM[P] = .032$, $GSD[P] = 2$, and $E[P] = .041$, such that only 5% of the penetration values exceed 0.1, in which case the respirator is deemed to have acceptable reliability.[3,4]

To begin, posit a geometric deterministic infectious dose distribution with parameter $\theta = 0.37$, and no cumulative dose effect. Equation (6) is used to compute values of $E[q \mid B = b]$ without and with respirator use; $b$ values up to 20 are used, because for $\theta = 0.37$, less than 0.01% of the cohort have an infectious dose greater than 20. Next, Equation (10) is used to compute the cohort infection risk over $M = 8$ periods without versus with respirator use. The latter cohort risks are, respectively, $CohR_8 = 0.95$ and 0.27. In the alternative, consider the same geometric deterministic infectious dose distribution with a cumulative dose effect over two days. Monte Carlo simulation is used to find $Pr[S_8 \geq B \mid B = b]$ for $b$ values up to 20, and Equation (12) is used to compute the cohort infection risk. Without versus with respirator use, after $M = 8$ periods the respective $CohR_8$ values are 1 and 0.32.

Fig. 1 depicts how infection risk accumulates over the exposure periods. Given no cumulative dose effect, Curve #1 corresponds to no respirator use; Curve #2 corresponds to respirator use. With a cumulative dose effect, Curve #3 corresponds to no respirator use; Curve #4 corresponds to respirator use. The figure demonstrates two ideas. First, a half-mask respirator is predicted to greatly reduce infection risk, although in this example the risk is still substantial. Second, where exposure periods are close in time, failing to account for a cumulative dose effect would lead to underestimating infection risk.

Interestingly, in this scenario the stochastic infectious dose model with $r = 0.37$ constant across all cohort members, $CohR_M = 1 - (E[q \mid r = 0.37])^M$, provides essentially the same risk estimates as does the deterministic model incorporating a cumulative dose effect.

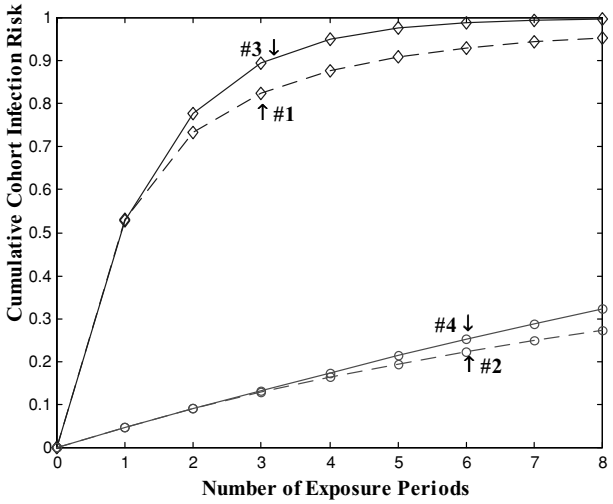


**Fig. 1.** Cumulative infection risk for up to eight exposure periods for a cohort whose infectious dose values for *C. immitis* spores follow a geometric distribution with parameter $\theta = 0.37$. Other model parameter values are described in the main text. Curve #1 is infection risk without a cumulative dose effect and without respirator use. Curve #2 is infection risk without a cumulative dose effect and with half-mask respirator use. Curve #3 is infection risk with a cumulative dose effect and without respirator use. Curve #4 is infection risk with a cumulative dose effect and with half-mask respirator use.

### 5.2. An Uncertainty Analysis

Because model parameters will be uncertain in almost every application, risk estimates will be uncertain. Traditional Monte Carlo simulation methods for uncertainty analysis permit placing probability bounds on the risk estimates and assessing the relative influence of different parameters. To measure the uncertainty importance of a given input (variable), we use the correlation coefficient of the sampled output values with the corresponding sample of values for the input:[28]

$$U_\rho(x_j, R) = \frac{\sum_{i=1}^{n}(x_{j,i} - \overline{x_j})(R_i - \bar{R})}{\sqrt{\sum_{i=1}^{n}(x_{j,i} - \overline{x_j})^2}\sqrt{\sum_{i=1}^{n}(R_i - \bar{R})^2}} \tag{14}$$

where $x_j$ denotes the $j^{th}$ uncertain input ($j = 1, \ldots, k$), $R_i$ is the risk estimate for the $i^{th}$ random value of $x_j$, and $\bar{R}$ is the mean value of R over $n$ values of $x_j$. The $U_\rho(x_j, R)$ values are inherently global measures of uncertainty importance over the joint probability distribution for the k inputs.

Consider a scenario analogous to the previous example. An eight-hour excavation is planned in a location suspected of *C. immitis* infection, for example, a site in an endemic region visibly contaminated with animal waste. The risk manager intends that each worker will use a respirator for four two-hour periods, with the aim of limiting the cohort infection risk to less than 0.05 with a high degree of confidence, say, 95%. The candidate respirator is a Type N100 full-face respirator. The construct to be applied is a deterministic infectious dose model with a geometric distribution and a cumulative dose effect.

Uncertainty in $\theta$ is modeled by a monotonically-increasing triangular density function on the interval [0.37, 1.0] with $E[\theta] = 0.79$. This $\theta$ range is consistent with the Table I dose-response data, and a triangular density function is posited because $\theta$ values closer to 1 have greater consistency with those data. Uncertainty in the future mean airborne spore concentration $E[C_A]$ is modeled by a log-normal distribution with a mean of 1 spore per $m^3$ and with GSD = 6 (a high degree of uncertainty corresponding to a percent CV $\approx$ 500%). Uncertainty in the mean respirator penetration value E[P] is modeled by a uniform distribution with a five-fold range centered on mean = 0.0041; the latter corresponds to the mean of a wearer's log-normal penetration distribution (with GSD = 2) such that the distribution conforms to acceptable respirator performance.[4] It is reasonable to treat $\theta$, $C_A$, and P as mutually independent. Other model inputs are the same as in the previous scenario for *C. immitis*.

Based on 50,000 random vectors of $\theta$, $E[C_A]$ and E[P], Fig. 2 shows the distribution of the output cumulative cohort infection risk, $CohR_4$. The best estimate is $\overline{CohR_4} = 0.033$, with 16% of values exceeding 0.05, and 5% exceeding 0.15. Although $\overline{CohR_4}$ nominally satisfies the intent that the cohort infection risk be less than 0.05, there is only 84% confidence that this constraint is met. To satisfy the desired degree of confidence, the risk manager would need to choose a respirator that permits less penetration, for example, a powered air-purifying respirator.[4] In this scenario, the three inputs exert the greatest effect on risk estimate uncertainty in the following order: mean spore concentration, mean respirator penetration value, and $\theta$. The respective sample correlation coefficients are 0.80, 0.15, and 0.069.

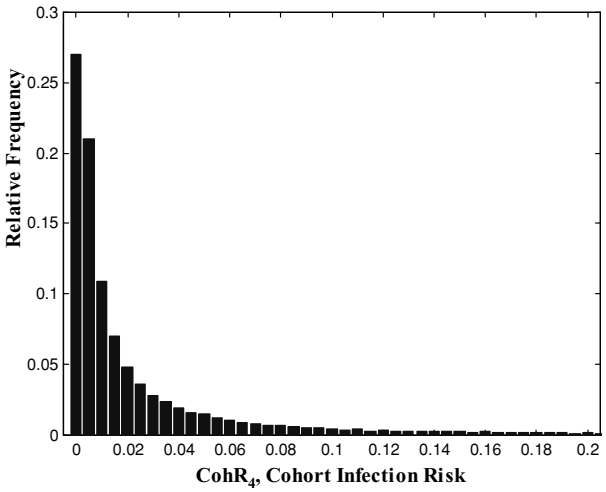


**Fig. 2.** The distribution of cohort risk estimates given four exposure periods and uncertainty in the values of the mean spore concentration, the mean respirator penetration value, and $\theta$. The posited distributions of these parameters, and the values of other model parameters, are described in the main text. The histogram is based on 50,000 random vectors of the uncertain inputs. The mean risk is $\overline{CohR_4} = 0.033$, with 16% of values exceeding 0.05.

## 6. DISCUSSION

The deficiency in the traditional Wells-Riley model is now readily explained. These investigators believed that a quantum unit could be used with the Poisson probability mass function to compute infection risk, where a quantum is that number of organisms equal to a deterministic infectious dose.[13,14] If $\mu$ denotes the expected number of quanta inhaled in a single exposure period, the Wells-Riley model equates infection risk with the Poisson probability of inhaling zero quanta: $R = 1 - \exp(-\mu)$. In contrast, our deterministic framework considers a person's infectious dose value $b$ and the expected number of organisms inhaled $\lambda$, and relates these to infection risk via Equation (5). The two constructs give the same risk estimate only if $b = 1$, in which case $\mu = \lambda$.

The Wells-Riley model leads to a positive error in estimating infection risk if $b > 1$, with the relative error increasing as $b$ increases. One result is that the model can fail to realistically predict risk reduction due to measures that reduce the inhaled pathogen concentration. To illustrate, assume that an individual's infectious dose value is $b = 10$, and that the expected number of organisms inhaled during an exposure period is $\lambda = 10$. Based on the Wells-Riley model with $\mu = 1$ (or $\lambda/b$), the infection risk is 0.63. Based on our Equation (5), the risk is 0.54. Next, consider that a control measure reduces the expected number of organisms inhaled to $\lambda = 1$, or one-tenth the original value. Based on the Wells-Riley model with $\mu = 0.1$, the infection risk is 0.095, which is still substantial.

Based on our Equation (5), the risk is $1.1 \times 10^{-7}$, which might be termed *de minimis*.

That said, in its original application to *M. tuberculosis* infection, the Wells-Riley model appears to be a good descriptor of infection risk because the biological evidence indicates that just one bacillus initiates infection.[13] The Wells-Riley model can be considered a special case of the deterministic model with $b = 1$ for all individuals and D a Poisson variable. In the alternative, it can be considered a special case of the stochastic one-hit model, $1 - (1 - \mathrm{r})^{\mathrm{D}}$, with $\mathrm{r} = 1$ for all individuals and D a Poisson variable.

Airborne pathogens have long been recognized as occupational hazards, but application of quantitative risk analysis to the workplace has been focused mainly on adverse health effects induced by chemical agents. Clearly, a quantitative framework for assessing airborne infection risk is an important supplement to the expert judgment that must be applied in selecting control measures such as respiratory protection. We remark that advances in the design of respiratory protection were stimulated by the use of poison gases and aerosols as chemical warfare agents during World War I; collateral benefits subsequently extended to workplace users.[3] We speculate that current attention to domestic bioterrorism will stimulate in like manner a more quantitative approach to assessing occupational infection risks and selecting control measures.

## APPENDIX

Groups of guinea pigs and *Macacus rhesus* monkeys were briefly exposed to spore aerosols. Based on the aerosol concentration, the duration of exposure, and the assumed breathing rate (150 mL/min for guinea pigs and 1200 mL/min for monkeys), the inhalation dose for each group was estimated. The endpoint scored was death of the animal. Table AI summarizes the data for guinea pigs exposed to spores carried by "4.5 $\mu$m [aerodynamic diameter] particles," and for monkeys exposed to "single-spore clouds."[11]

**Table AI**

| Dose (# Spores Inhaled) | # Dead/# Exposed |
|---|---|
| Guinea Pigs | |
| 18,750 | 6/40 |
| 34,650 | 18/40 |
| 40,350 | 17/40 |
| 56,100 | 13/40 |
| 57,750 | 22/40 |
| 89,550 | 22/40 |
| 153,750 | 28/40 |
| Monkeys | |
| 35,160 | 1/8 |
| 38,520 | 4/8 |
| 54,360 | 5/8 |
| 68,760 | 6/8 |
| 77,760 | 5/8 |
| 80,400 | 3/8 |
| 120,000 | 8/8 |
| 150,000 | 7/8 |
| 199,200 | 8/8 |

The fit of four models was investigated. The first is a single-parameter stochastic model (referred to as the *simple binomial* model):

$$\pi(\mathrm{D}) = 1 - (1 - r)^{\mathrm{D}} \qquad \text{(A1)}$$

where D is the dose (number) of organisms received, r is the probability that any single organism infects the host, and $\pi(\mathrm{D})$ is the host's probability of becoming infected. For simplicity, it was assumed that infection always led to death. In this model, all organisms have the same probability of causing infection, which does not vary between hosts (all hosts have equal susceptibility), and the probability of any particular organism causing infection is independent of all others. The only parameter to be estimated is $\theta = \log(\mathrm{r})$.

The second model is an extension of the first. For a given host, Equation (A1) applies, but the value of r varies between hosts according to a beta distribution (referred to as the *beta binomial* model). Specifically, the model is:

$$\pi_{\mathrm{i}}(\mathrm{D}) = 1 - (1 - \mathrm{r_i})^{\mathrm{D}} \qquad \text{(A2)}$$

where $\mathrm{r_i}$ is the probability of infection per organism in the $i^{th}$ host, and $\mathrm{r_i}$ follows a standard beta distribution across hosts with parameters $\alpha$ and $\beta$. This is a two-parameter model with $\theta_1 = \ln(\alpha)$ and $\theta_2 = \ln(\beta)$.

The third model assumes that the dose required to infect a host is deterministic, but the deterministic doses vary across hosts according to a log-normal distribution (referred to as the *log-normal* model). The probability that a randomly selected host is infected by a dose D is the probability that the host's infection threshold B is less than D:

$$\pi(\mathrm{D}) = \Phi\left(z = \frac{\log(\mathrm{D}) - \mu_\ell}{\sigma_\ell}\right) \qquad \text{(A3)}$$

where B follows a log-normal distribution with parameters $\theta_1 = \mu_\ell$ and $\theta_2 = \sigma_\ell$, and $\Phi(\mathrm{z})$ is the cumulative standard normal distribution evaluated at the argument $z$. Note that $\mu_\ell$ is the logarithm of the geometric mean, and $\sigma_\ell$ is the logarithm of the geometric standard deviation.

**Table AII**

| Model | Guinea Pig Data $\theta_1, \theta_2$ | Guinea Pig Data AIC | Monkey Data $\theta_1, \theta_2$ | Monkey Data AIC |
|---|---|---|---|---|
| Simple binomial | −11.5, NA | 366.5 | −11.1, NA | 77.1 |
| Beta binomial | 4.0, 15.5 | 368.5 | 4.1, 15.1 | 77.4 |
| Log-normal | 11.1, 1.6 | 364.4 | 10.9, 0.72 | 78.2 |
| Logistic | $-1.1, 1.4 \times 10^{-5}$ | 368.4 | $-1.7, 3.0 \times 10^{-5}$ | 80.0 |

The fourth model is an empirical, logistic dose-response construct (referred to as the *logistic* model) that treats the probability of infection as a function of dose:

$$\pi(\mathrm{D}) = \frac{1}{1 + \exp(-(\alpha + \beta \cdot D))} \tag{A4}$$

This is also a two-parameter model with $\theta_1 = \alpha$ and $\theta_2 = \beta$.

The four models were fit using maximum likelihood to the guinea pig and monkey data. The evaluation of fit was done using Aikaki's information criterion.[29] This statistic measures the relative fit of each model to the data, while penalizing the fit for the number of parameters. The statistic does not represent a definitive test of the appropriateness of each model. Rather, it is a measure of the relative evidence from the data for each of the four models. A smaller AIC value represents a better relative penalized fit. The results of the model fitting are summarized in Table AII.

Roughly, a difference in the AIC of two or more is considered a potentially superior fit. For the guinea pig data, the deterministic log-normal dose model appears to fit the data best. For the monkey data, the first three models are essentially equivalent in their fit, with the logistic model appearing to be a less compelling fit to the data.

## ACKNOWLEDGMENTS

The authors thank Dr. Robert C. Spear for his guidance in the analysis. This work was funded by NIOSH/CDC Award No. 5 R01 OH03570. The ideas expressed are solely the authors' and do not necessarily represent the official views of the funding agency.

## REFERENCES

1. Occupational Safety and Health Administration. (1998). 29 CFR, Parts 9110 and 1926, "Respiratory Protection, Final Rule." *Federal Register, 63*, 1152–1300.
2. Pritchard, J. A. (1976). *A guide to industrial respiratory protection*. Cincinnati, OH: National Institute for Occupational Safety and Health, U.S. Department of Health, Education and Welfare.
3. National Institute for Occupational Safety and Health. (1987). *Respirator decision logic*. Cincinnati, OH: U.S. Department of Health, Education and Welfare.
4. American National Standards Institute Z88.2 Committee. (1992). *American national standard for respiratory protection*. New York: American National Standards Institute.
5. Khan, A. S., Levitt, A. M., & Sage, M. J. (2000). Biological and chemical terrorism: Strategic plan for preparedness and response. *Morbidity and Mortality Weekly Report, 49*(RR-4), 1–14.
6. Centers for Disease Control and Prevention. (1993). Hantavirus infection—Southwestern United States: Interim recommendations for risk reduction. *Morbidity and Mortality Weekly Report, 42*(RR-11), 1–12.
7. Centers for Disease Control and Prevention. (1994). Guidelines for preventing the transmission of *Mycobacterum tuberculosis* in health-care facilities, 1994. *Morbidity and Mortality Weekly Report, 43*(RR-3), 1–32.
8. McCullough, N. V., & Brosseau, L. M. (1999). Selecting respirators for control of worker exposure to infectious aerosols. *Infection Control and Hospital Epidemiology, 20*, 136–144.
9. Nicas, M., Neuhaus, J., & Spear, R. C. (2000). Risk-based selection of respirators against infectious aerosols: Application to anthrax spores. *Journal of Occupational and Environmental Medicine, 42*, 737–748.
10. Defense Intelligence Agency. (1986). *Soviet biological warfare threat.* U.S. Department of Defense, DST-1610F-057-86. Washington, DC: U.S. Department of Defense.
11. Druett, H. A., Henderson, D. W., Packman, L., & Peacock, S. (1953). Studies on respiratory infection. I. The influence of particle size on respiratory infection with anthrax spores. *Journal of Hygiene, 51*, 359–371.
12. Pappagianis, D. (1988). Epidemiology of coccidioidomycosis. In M. R. McGinnis (Ed.), *Current topics in medical mycology* (Vol. 2, pp. 199–238). New York: Springer-Verlag.
13. Wells, W. F. (1955). *Airborne contagion and air hygiene* (pp. 104–141). Cambridge, MA: Harvard University Press.
14. Riley, R. L., & O'Grady, F. O. (1961). *Airborne infection: Transmission and control* (pp. 95–125). New York: MacMillan Company.
15. Haas, C. N. (1983). Estimation of risk due to low doses of microorganisms: A comparison of alternative methodologies. *American Journal of Epidemiology, 118*, 573–582.
16. Regli, S., Rose, J. B., Haas, C. N., & Gerba, C. P. (1991). Modeling the risk from Giardia and viruses in drinking water. *Journal of the American Water Works Association, 83*, 76–84.
17. Teunis, P. F. M., van der Heijden, O. G., van der Giessen, J. W. B., & Havelaar, A. H. (1996). *The dose-response relation in human volunteers for gastro-intestinal pathogens*. National Institute of Public Health and Environment, report no. 284550002. Bilthoven, the Netherlands: National Institute of Public Health and Environment.
18. Busch, K. A., & Leidel, N. A. (1988). Statistical models for occupational exposure measurements and decision making. In American Conference of Governmental Industrial Hygienists, *Advances in air sampling*. Chelsea, MI: Lewis Publishers.
19. Galvin, K., Selvin, S., & Spear, R. C. (1990). Variability in protection afforded by half-mask respirators against styrene exposure in the field. *American Industrial Hygiene Association Journal, 51*, 625–631.
20. Nicas, M., & Spear, R. C. (1992). A probability model for assessing exposure among respirator wearers: Part I—Description of the model. *American Industrial Hygiene Association Journal, 53*, 411–418.

21. Harpe, G. J., & Morton, J. D. (1953). The respiratory retention of bacterial aerosols: Experiments with radioactive spores. *Journal of Hygiene, 51*, 372–385.
22. Stevens, D. A. (1995). Coccidioidomycosis. *New England Journal of Medicine, 332*, 1077–1082.
23. Kirkland, T. N., & Fierer, J. (1996). Coccidioidomycosis: A reemerging infectious disease. *Emerging Infectious Diseases, 2*, 192–199.
24. Division of Communicable Disease Control. (2000). *California morbidity table of selected reportable diseases, 1998*. California Department of Health Services. Available at http://www.dhs.ca.gov/ps/dcdc/html/cdtables.htm.
25. Blundell, G. P., Castleberry, M. W., Lowe, E. P., & Converse, J. L. (1961). The pathology of *Coccidioides immitis* in the Macca Mulatta. *American Journal of Pathology, 39*, 613–630.
26. Converse, J. L., & Reed, R. E. (1966). Experimental epidemiology of coccidioidomycosis. *Bacteriological Reviews, 30*, 678–695.
27. Converse, J. L., Lowe, E. P., Castleberry, M. W., Blundell, G. P., & Besemer, A. R. (1962). Pathogenesis of *Coccidioides immitis* in monkeys. *Journal of Bacteriology, 83*, 871–878.
28. Morgan, M. G., & Henrion, M. (1998). *Uncertainty, a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge, UK: Cambridge University Press.
29. Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control, AU-19*, 716–722.