Mathematically Assessing the Consequences of Food Terrorism Scenarios

Y. Liu and L.M. Wein

ABSTRACT: We derive mathematical expressions for the mean number of casualties resulting from a deliberate release of a biological or chemical agent into a food supply chain. Our analysis first computes the amount of contaminated food as a function of the network topology and vessel sizes in the food processing plant. A probabilistic analysis, in which each potential consumer of contaminated food has his own random purchase time, infectious dose, and incubation period, determines the number of people who consume enough tainted food to get infected or poisoned before the attack is detected and food consumption is halted. These simple formulas can be used by the U.S. government and the food industry to develop a rough-cut prioritization of the threats from food terrorism, which would be a 1st step toward the allocation of appropriate prevention and mitigation resources.

Keywords: bioterrorism, homeland security, mathematical modeling.

Introduction

It is not clear whether botulinum toxin in milk is the unique catastrophic scenario in food terrorism (Wein and Liu 2005) or whether it is merely the tip of the iceberg. Given the wide array of biological or chemical agents that could be used in a food attack and the immense variety of food types that could be targeted, it is not practical to construct and analyze from scratch a detailed mathematical model for each agent and food type. Nonetheless, as noted in Recommendation nr 14 in a recent report by the Office of Inspector General, Dept. of Homeland Security (2007), it is vitally important to quickly assess all of the agent-food type scenarios so that resources for prevention and mitigation can be allocated rationally. Toward this end, we analyze a generic model of an attack on the food supply and derive formulas for the mean number of casualties in terms of the key characteristics of the biological or chemical agent and the food supply chain. Given the intended use of our model to aid in a rough-cut ranking of various agent–food type combinations, we sacrifice some realism and accuracy to preserve some parsimony and ease of use.

The Model

We begin with a brief overview. The model tracks contaminated food as it moves through the various stages of storage and processing within the food processing facility. The model (1) allows food to move continually through the facility or to be processed in discrete batches; (2) permits the volume of food to expand or contact at each stage; (3) allows for forking (that is, disassembly) and assembly operations within the facility; and (4) assumes that a fraction of the contaminating agent survives storage, processing, transportation, and final food preparation. In the downstream portion of the supply chain, each unit of contaminated food (for example, a gallon of milk) is purchased by a small group (for example, a family) of potential consumers after a random distribution delay, and then this unit of food is consumed by each consumer in the small group at a constant rate over a finite amount of time. We use a dose–response curve to determine if and when each consumer becomes infected or poisoned. Each infected or poisoned consumer displays symptoms after a random incubation period, the attack is detected after a specified number of consumers become symptomatic, and all food consumption is halted after a further specified delay.

We now turn to the detailed model formulation. The raw data for the model are the values for the parameters described in Table 1. The food processing facility is modeled as a set of N stages of storage or production. Each stage uses either a continuous flow process or a batch process (Schmenner 1993), the distinction being that food continually moves through the system as it is being processed in the former system, and is processed in large batches in discrete steps in the latter system. The only difference in the analysis stems from whether stage 1 is batch flow or continuous. In the model, a deliberate biological or chemical release of size Q occurs at stage 1, and if a downstream stage in a facility is the obvious release location, then the upstream portion of the facility can be ignored and this release location can be viewed as stage 1. Throughout the supply chain, the agent is assumed to be uniformly mixed into the food. While this assumption is realistic for many liquids (for example, milk silos, which would be stage 1 in a typical milk attack [Wein and Liu 2005], have agitators that continually mix the milk), it may not be realistic for solids (and may lead us to overestimate the magnitude of an attack involving solid foods). Each stage $i = 1, \ldots, N$ is partially characterized by 3 variables: the number of containers (these can be for storage or processing) (n_i); the volume transformation factor from stage i to stage $i+1$ (θ_i), where $\theta_i < 1$ represents, for example, concentration or evaporation, and $\theta_i > 1$ represents, for example, expansion or addition of material; and the type of operational flow from stage i to stage $i+1$ (o_i), where $o_i = d$ for dedicated, $o_i = f$ for forked, $o_i = a$ for assembly, and $o_i = m$ for full mixing (Figure 1). For simplicity, we do not consider partial mixing, and we assume that $\max\{\frac{n_i}{n_{i+1}}, \frac{n_{i+1}}{n_i}\}$ is an integer for $i = 1, \ldots, N-1$. In the main text, we assume that the agent of concern is essentially inert, that is, does not undergo microbial growth, during processing, distribution, and consumption; however, at the end of the Appendix, we briefly outline how to generalize our analysis to allow for growth of microbrial agents throughout the supply

MS 20080101 Submitted 2/7/2008, Accepted 5/13/2008. *Author Liu is with Dept. of Systems Engineering and Operations Research, George Mason Univ., Fairfax, VA 22030. Author Wein is with Graduate School of Business, Stanford Univ., Stanford, CA 94305-5015. Direct inquiries to author Wein (E-mail: lwein@stanford.edu)*

doi: 10.1111/j.1750-3841.2008.00874.x

chain via, for example, the Gompertz or Baranyi-Roberts (Baranyi and Roberts 1994) growth model.

In addition to these 3 variables, we let λ_1 be the total arrival rate of food to stage 1, and λ_N, which equals $(\prod_{j=1}^{N} \theta_j)\lambda_1$, be the rate at which food exits stage N. We let s_i denote the size of each of the n_i containers in stage i, and impose a balanced capacity assumption, $n_i\,\theta_i\,s_i = n_{i+1}s_{i+1}$, which allows us to perform all calculations without using $s_2, \ldots, s_N$. If stage 1 is a continuous flow process then we define w_1 to be the time between successive cleanings of stage 1 equipment (for example, silos). In this case, the amount of contaminated food depends on when during the cleaning cycle the attack occurs (Wein and Liu 2005). For mathematical simplicity, in this case we compute the mean number of casualties resulting from the mean amount of contaminated food, rather than taking the expectation over the amount of contaminated food; the loss of accuracy due to this assumption is investigated later. Also, if stage 1 is a continuous flow process, the contaminated food will have different concentration levels of agent because the concentration level at stage 1 drops exponentially just after the instantaneous point release due to the simultaneous arrival of uncontaminated food and the outflow of contaminated food (Wein and Liu 2005). To simplify our analysis, we assume that all of the contaminated food has the same concentration level of agent by the time it exits stage N. Although this is an approximation for a continuous flow system, the concentration levels do become much more homogeneous if food is assembled into large containers downstream (Wein and Liu 2005), as is typical in many food processes.

Table 1—Model parameters.

N	Number of stages of storage and processing
Q	Amount of biological or chemical agent released
n_i	Number of containers at stage i for $i = 1, \ldots, N$
θ_i	Volume transformation factor from stage i to stage $i+1$ for $i = 1, \ldots, N$
o_i	Type of operational flow from stage i to stage $i+1$ for $i = 1, \ldots, N-1$
λ_1	Arrival rate of food to stage 1
w_1	Time between successive cleanings at stage 1
s_i	Size of each container at stage i for $i = 1, \ldots, N$
α	Fraction of agent that survives storage, processing, and transportation
n	Consumers per unit of food
δ	Minimum speed of distribution channel
e^{μ_d}	Median speed of distribution channel
e^{σ_d}	Dispersal factor of speed of distribution channel
ID_{50}	Median infectious dose
β	Probit slope of the dose–response function
e^{μ_s}	Median incubation period
e^{σ_s}	Dispersal factor of incubation period
c	Food consumption rate
k	Number of symptomatics until detection
Δ	Time from detection to halting of consumption

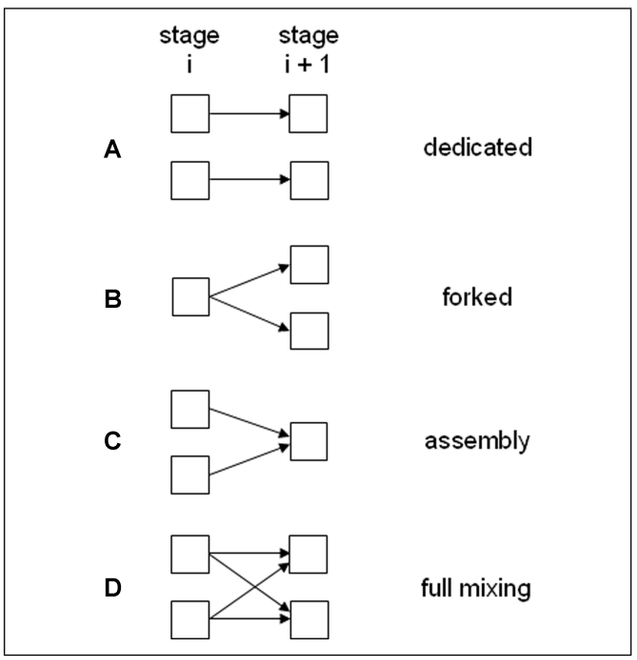

Figure 1—The 4 types of operational flow from stage i to stage $i+1$, illustrated for the special case in which $\max\{n_i, n_{i+1}\} = 2$. (A) In dedicated flow, $n_i = n_{i+1}$ and each of the n_i containers feeds a corresponding container in stage $i+1$. (B) In forked flow, $\frac{n_{i+1}}{n_i}$ is an integer greater than 1, and the contents of each of the n_i containers at stage i are split evenly among $\frac{n_{i+1}}{n_i}$ containers at stage $i+1$. (C) In assembly flow, $\frac{n_i}{n_{i+1}}$ is an integer greater than 1, and the contents of each set of $\frac{n_i}{n_{i+1}}$ containers at stage i are assembled into a single container at stage $i+1$. (D) In full mixing, all contents from all n_i containers at stage i are split evenly among all n_{i+1} containers at stage $i+1$.

A fraction α of the biological or chemical agent survives during storage, processing, transport, and final food preparation. There are n potential consumers per unit (for example, gallon or pounds) of contaminated food; that is, if the attack goes undetected then each person consumes $1/n$ of a unit of contaminated food. Each potential food consumer in our model has an associated independent and identically distributed (i.i.d.) random variable that represents the time of purchase (that is, the amount of time from when the 1st contaminated food leaves the production facility to the time when the potential consumer purchases his contaminated food). We decompose this random variable into 2 pieces, with the 1st piece being the time interval from when the 1st contaminated food leaves the facility until the time when the contaminated food consumed by a random consumer leaves the facility. This time interval is a uniform random variable with a mean and standard deviation that are calculated from primitive model parameters by assuming that contaminated food leaves the facility at a constant rate, which is the output rate λ_N times the fraction of food at stage N that is contaminated. The 2nd piece, which we refer to as the speed of the distribution channel, is the interval from when the contaminated food of a random consumer leaves the facility until it is purchased. Because the distribution time has a minimum positive value and is typically skewed to the right, we assume that the speed of the distribution channel is the sum of a constant δ and a lognormal random variable with median e^{μ_d} and dispersal factor e^{σ_d} (Table 1).

Upon purchase, we assume that each person consumes contaminated food at the continuous rate c (hence, we are ignoring the food preparation time, which is typically much less than the time it takes for the food to travel through the distribution channel), so that his allotment of $1/n$ food units is consumed in $1/c$ time units. Note that although the shelf life of a product is highly tied to the speed of the production and distribution system, the amount of time the product is offered for sale, and the period over which the product is consumed, we do not explicitly incorporate the product shelf life into the model. Rather, certain quantities, such as the consumption rate and the mean and standard deviation of the distribution system, are primitive model parameters, with the understanding that these quantities are partly driven by the product shelf life.

We employ a probit dose–response function, so that a person who consumes y food units is infected (we use "infected" although some agents cause poisoning) with probability $\Phi[\beta \log_{10}(\frac{y}{\mathrm{ID}_{50}})]$, where ID_{50} is the dose that infects half of the population and β is the probit slope (Wein and Liu 2005); as is typical in the infectious disease literature (for example, Wells 1955), we assume that the dose is a continuous variable. Equivalently, each potential consumer has an i.i.d. uniform [0, 1] random variable u and gets infected if he consumes no less than his infectious dose, which is $Y = \mathrm{ID}_{50}\, 10^{\Phi^{-1}(u)/\beta}$. A person becomes infected at the moment he has consumed his infectious dose Y. He then incurs a log normally distributed incubation period with median e^{μ_s} and dispersal factor e^{σ_s}, which is the time delay between the onset of infection and the onset of symptoms. We assume that the attack is detected when the k^{th} person develops symptoms (hence, the incubation period actually includes the time it takes for a symptomatic person to report to a medical facility so that he can be diagnosed). All food consumption is halted after a further delay Δ.

Detection and consumption halting are exceedingly complex and we have taken a rather simple cut at them, which we believe is consistent with the goals of our model (both in terms of accuracy and ease of use). The actual detection time will depend on a variety of issues that are not explicitly modeled, including the amount of information on the food's container (for example, cartons of milk contain information about which shift of which processing plant it originated from), the spatial distribution of the food, and the quality of the medical surveillance network (for example, whether spatially distributed symptomatic patient data are shared). Similarly, the parameter Δ depends on a variety of unmodeled variables, such as the accuracy and speed of communication from epidemiologists to decision makers, the decisiveness of the decision makers, and the efficacy of the communication approaches used to inform the public about the attack.

Results

In the Appendix, we derive an approximate (the approximation is restricted to our estimate of the time at which consumption is halted in Eq. 8) but reasonably accurate estimate for the mean number of casualties (that is, number of people poisoned or infected), denoted by $E[I]$, and assess its accuracy against simulation results from the exact model. We present the results as a 2-step procedure, where we first compute the mean amount of contaminated food, denoted by H, by a simple iterative algorithm. For stage $i = 1, \ldots, N$, let m_i be the number of the n_i containers that have contaminated food, and let h_i be the amount of contaminated food. Starting with the initial conditions

$$m_1 = 1 \quad \text{and} \quad h_1 = \begin{cases} s_1 & \text{if stage 1 is batch} \\ \dfrac{\lambda_1 w_1}{2n_1} & \text{if stage 1 is continuous flow} \end{cases} \tag{1}$$

we iteratively solve

$$m_{i+1} = \begin{cases} m_i & \text{if } o_i = d \\ \dfrac{n_{i+1} m_i}{n_i} & \text{if } o_i = f \\ \left\lceil \dfrac{n_{i+1} m_i}{n_i} \right\rceil & \text{if } o_i = a; \\ n_{i+1} & \text{if } o_i = m \end{cases} \tag{2}$$

$$h_{i+1} = \begin{cases} \theta_i h_i & \text{if } o_i = d \\ \theta_i h_i & \text{if } o_i = f \\ \dfrac{m_{i+1} n_i}{m_i n_{i+1}} h_i \theta_i & \text{if } o_i = a \\ n_1 h_1 \left(\prod_{j=1}^{i} \theta_j \right) & \text{if } o_i = m \end{cases} \tag{3}$$

for $i = 1, \ldots, N-1$, where $\lceil x \rceil$ is the smallest integer greater than or equal to x, and set

$$H = h_N \tag{4}$$

In terms of H in Eq. 4, the mean number of casualties is

$$E[I] = nH\Phi\left(\beta \log_{10}\left(\frac{\frac{\alpha Q}{nH}}{\mathrm{ID}_{50}}\right)\right) \quad \text{if } nH\Phi\left(\beta \log_{10}\left(\frac{\frac{\alpha Q}{nH}}{\mathrm{ID}_{50}}\right)\right) < k \tag{5}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. In this case, the attack is too small to be detected. If $nH\Phi(\beta \log_{10}(\frac{\alpha Q/nH}{\mathrm{ID}_{50}})) \geq k$, then the mean number of casualties is

$$E(I) = nH \int_0^\infty \int_0^{\frac{n_N H}{m_N \lambda_N}} \Phi\left(\beta \log_{10}\left(\frac{C_j}{\mathrm{ID}_{50}}\right)\right) \times \frac{m_N \lambda_N}{n_N H} \frac{1}{x_2 \sigma \sqrt{2\pi}} e^{-\frac{(\ln x_2 - \mu_d)^2}{2\sigma_d^2}} dx_1 dx_2 \tag{6}$$

where

$$C_j = \frac{\alpha Qc \max(0, \min(c^{-1}, \tau - (x_1 + \delta + x_2)))}{nH} \tag{7}$$

and

$$\tau = \exp\left(\mu_w + \Phi^{-1}\left(\frac{k}{nH}\right)\sigma_w\right) + \Delta \tag{8}$$

is the time at which consumption is halted, where μ_w and σ_w are solved from

$$e^{\mu_w + \frac{\sigma_w^2}{2}} = \frac{n_N H}{2 m_N \lambda_N} + \delta + e^{\mu_d + \frac{\sigma_d^2}{2}} + \frac{nH\mu_i}{\alpha Qc} + e^{\mu_s + \frac{\sigma_s^2}{2}} \tag{9}$$

$$\left(e^{\sigma_w^2} - 1\right) e^{2\mu_w + \sigma_w^2} = \frac{n_N^2 H^2}{12 m_N^2 \lambda_N^2} + \left(e^{\sigma_d^2} - 1\right) e^{2\mu_d + \sigma_d^2} + \left(\frac{nH\sigma_i}{\alpha Qc}\right)^2 + \left(e^{\sigma_s^2} - 1\right) e^{2\mu_s + \sigma_s^2} \tag{10}$$

Equation 5 also provides a simple estimate for the mean number of casualties in an attack of any size under the pessimistic assumption that the attack goes undetected. Although this estimate requires less data and is simpler to compute than Eq. 6, it is also cruder and fails to account for how the consumption rate and the probability distributions of the purchase time and incubation period interact to determine the efficacy of halting consumption.

To assess the accuracy of Eq. 6, we compare its estimates to the exact simulated values (the Monte Carlo simulation generates random purchase times, infectious doses and incubation periods for all nH potential consumers to determine the number of casualties,

and then repeats this procedure many times until the 95% confidence interval half-width for the number of casualties is < 1% of the mean number of casualties) for various release sizes, using parameter values from a botulinum toxin in milk scenario in Wein and Liu (2005) study; readers are referred to Wein and Liu (2005) for a discussion and derivation of these parameter values. In particular, by Eq. 1, we set $H = h_1 = \frac{\lambda_1 w_1}{2n_1} = \frac{(640000 \text{ gallons/day})(72 \text{ h})}{2(8)} = 120000$ gallons because there are 8 silos, and set $\lambda_N = \lambda_1$, $m_N = 1$, and $n_N = 8$ (that is, perfect yield and dedicated processing lines). We use the values $\alpha = 0.316$ (this value of α was derived in Wein and Liu [2005] from the inactivation rate of botulinum toxin in canned corn, which has a similar pH to milk and is based on an outdated heat pasteurization process; the heat pasteurization process in the U.S. dairy industry has been intensified in recent years [Alberts 2005], and the true value of α, which is proprietary, may be much different than 0.316), $n = 4$/gallon, $\delta = 48$ h, $e^{\mu_d} = 9.55$ h, $e^{\sigma_d} = 2.40$, $ID_{50} = 1$ μg, $\beta = 4.34$, $e^{\mu_s} = 46.99$ h, $e^{\sigma_s} = 1.23$, $c^{-1} = 84$ h, $k = 100$, and $\Delta = 24$ h, and then compute $E[I]$ for various values of Q. The numerical results (Table 2) reveal that Eq. 6 is sufficiently accurate for the purpose of rough-cut prioritization: within 1% of the exact value for small (≤ 0.5 g) attacks, and within $\approx 10\%$ for larger (> 1 g) attacks. We stress that the goal of the comparison is not to assess the magnitude of a botulinum in milk attack (which is not possible without an accurate value of α), but rather to assess the accuracy of Eq. 6.

Table 2—Simulated compared with approximate mean casualties resulting from the mean amount of contaminated food. For various release sizes (Q), we compare the simulated mean number of casualties to the mean casualties estimated by Eq. 6.

Release size	Mean casualties simulated	Mean casualties Eq. 6
0.18 g	13.29	13.96
0.2 g	31.83	31.84
0.3 g	526.8	526.9
0.5 g	7852	7882
0.8 g	37473	39931
0.9 g	49473	53420
1 g	61728	67104
10 g	417430	367840
100 g	446500	413400
1 kg	446780	414530
10 kg	446770	414610

The simulated values in Table 2 are the mean number f casualties resulting from the mean amount of contaminated food. However, the amount of contaminated food is itself a uniform random variable between 0 and $\frac{\lambda_1 w_1}{n_1}$. Taking into account the random amount of contaminated food and the fact that the number of casualties for a given amount of contaminated food is actually a binomial random variable, we use Monte Carlo simulation to create a histogram (of $\approx 10^5$ instances) for the number of casualties when the release size is 1 g (Figure 2). The variation in the number of casualties is dictated nearly entirely by the random amount of contaminated food, and so the upper bound on the number of casualties is roughly twice the mean value. The u-shape of the histogram in Figure 2 can be explained by the plot of casualties compared with amount of contaminated food in Figure 3: the right peak in Figure 2 is due to the peak in Figure 3, and the left peak in Figure 2 is due to the skewed right tail in Figure 3. In addition, the expected value of the number of casualties in Figure 2 is 57391, compared with the value of 61728 in Table 2, which is the mean number of casualties resulting from the mean amount contaminated food. Hence, our qualitative results are not changed by considering only the mean amount of contaminated food in Eq. 6.

Discussion

A probabilistic risk assessment typically includes a threat analysis (what is the likelihood of an attack), a vulnerability analysis (what is the likelihood the attack will be successful), and a consequence analysis (what is the damage from a successful attack). Because the U.S. government is likely to have very little reliable detailed data about terrorists' intentions and capabilities that would inform a threat analysis, and because a determined and resourceful terrorist group would be capable of defeating any security mechanisms that might be put in place at a food processing facility, a probabilistic risk analysis of a food terrorist attack is almost entirely dictated by a consequence analysis, such as that performed here.

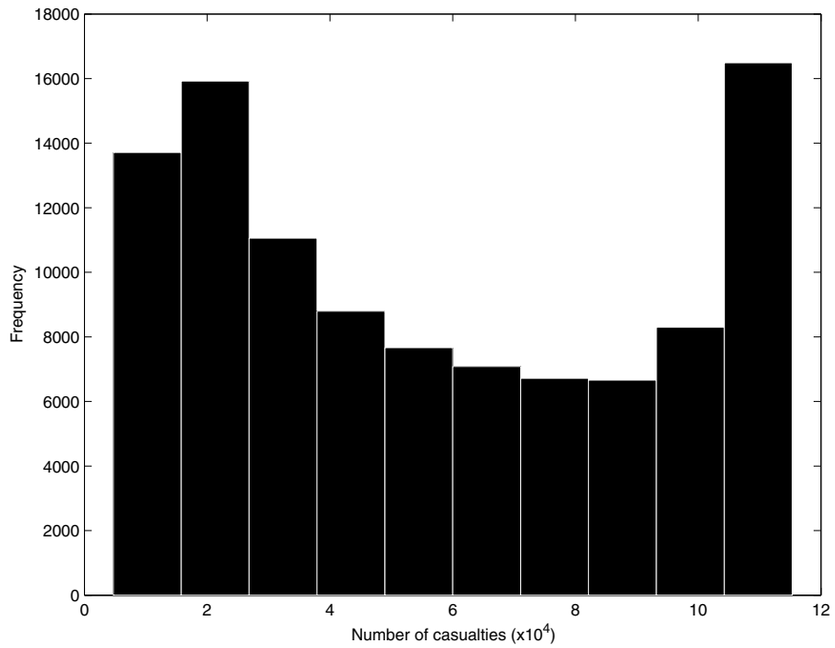

Figure 2—For a release size of 1 g, the simulated histogram (using $\approx 10^5$ instances) of the number of casualties, accounting for the randomness in the amount of contaminated food.

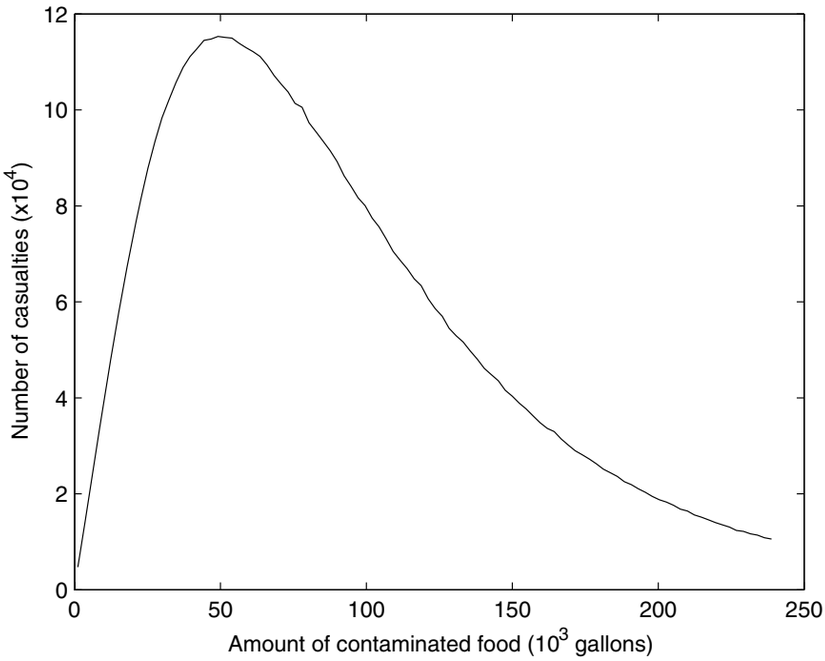

Figure 3—For a release size of 1 g, the simulated number of casualties compared with the amount of contaminated food.

Section 112(r) of the Clean Air Act forced every chemical facility with large quantities of hazardous chemicals to perform an offsite consequence analysis, which estimates the number of people endangered by a chemical accident. These analyses, the majority of which were based on methodology proposed by the Chemical Emergency Preparedness and Prevention Office, U.S. EPA, are providing the basis for post-9/11 security upgrade requirements and alternative technology assessments. Because of national security concerns, the results of these analyses are not in the public domain (Dept. of Justice 2000). In a similar way, new laws may be needed to require food production facilities to perform consequence analyses. In our view, because of the enormous number of agent–food combinations, and because only a small fraction of these combinations is likely to result in a large number of casualties, it is appropriate to use a rather crude and easy-to-use method (that is, an idealized model with parsimonious data requirements) to initially prioritize all combinations, and then to follow up with a more detailed analysis of the minority of combinations that are potentially dangerous. In this context, Eq. 1 to 10 can be viewed as a 1st draft (for example, the analysis could be elaborated with a mortality rate or some other measure that quantifies the severity of the casualties associated with a particular biological or chemical agent) of the methodology required for the initial prioritization.

However, a more detailed analysis of high-priority combinations might generalize our model in several directions, including an investigation of the extent of non uniform mixing (particularly for solid foods), a more elaborate detection model that incorporates the spatiotemporal clustering of cases, the amount of information sharing, the quality of the surveillance system, and the uniqueness of symptoms (although the data requirements for such a study would be more onerous than for our model), more realistic time lag distributions, and more detailed outcome measures, such as the distribution of the number of casualties.

The values for the parameter α, which is the fraction of agent that survives storage, production, and transportation, are not typically available outside of the food industry (at least for biological and chemical agents that do not naturally arise in food safety—as opposed to food security scenarios), and indeed may even require experiments to estimate them. The ComBase Initiative, which is an international collaboration that created the ComBase Database (Baranyi and Tamplin 2004) and the microbial models in the ComBase Predictor (Combase 2008), has publicly available software that computes the time-dependent concentration of microbial agents in food using a growth (for example, Gompertz or Baranyi-Roberts) or survival function with parameters that are polynomial functions of quantities such as temperature, pH, and NaCl concentration; this could be very helpful in computing α for various agent–food combinations. As with the results of the chemical consequence analyses, it is likely that the U.S. government would not want the detailed results of a food consequence analysis to be made public. Consequently, we do not derive numerical estimates for various agent–foodtype scenarios in this article, but we urge the U.S. government to do so, either by enacting new laws or via its new SPPA Initiative (2007). Our software is available for free to U.S. government and food industry employees upon request.

Finally, while our focus is on food security, our results also apply to an accidental contamination of food, particularly if the model is extended to growth of microbial agents as outlined in the Appendix. For example, these results could help the U.S. FDA choose the types of foods and accidental hazards that should be subject to the Hazard Analysis and Critical Control Point (HACCP) approach to managing food safety (FDA 2007).

Acknowledgment

This study was partially supported by the Center for Social Innovation, Graduate School of Business, Stanford Univ.

References

Alberts B. 2005. Modeling attacks on the food supply. PNAS 102:9737–8.
Balkema AA, De Haan L. 1978. Limit distributions for order statistics, II. Theory Prob Applic 23:341–58.
Baranyi J, Roberts TA. 1994. A dynamic approach to predicting bacterial growth in food. Int J Food Microbiol 23:277–94.
Baranyi J, Tamplin ML. 2004. ComBase: a common database on microbial responses to food environments. J Food Prot 67:1967–71.

Buchanan RL. 1991. Using spreadsheet software for predictive microbiology applications. J Food Saf 11:123–34.
Chemical Emergency Preparedness and Prevention Office, U.S. EPA. 1999. Risk management program guide for offsite consequence analysis. EPA-550-B-99-009. Washington, D.C.: Environmental Protection Agency.
Combase. 2008. Combase—common database for predictive microbiology. Available from: http://ww.combase.cc/. Accessed Apr 1, 2008.
Dept. of Justice. 2000. Assessment of the increased risk of terrorist or other criminal activity associated with posting off-site consequence analysis information on the Internet. Washington, D.C.: Dept. of Justice.
[FDA] Food and Drug Administration. 2007. HACCP: a state-of-the-art approach to food safety. Available from: http://www.cfsan.fda.gov/~lrd/bghaccp.html. Accessed Jun 12, 2007.
Gross D, Harris CM. 1985. Fundamentals of queueing theory. 2nd ed. New York: John Wiley and Sons.
Office of Inspector General, Dept. of Homeland Security. 2007. The Department of Homeland Security's role in food defense and critical infrastructure protection. Report OIG-07-33, Washington, D.C.
Schmenner RW. 1993. Production/operations management. 5th ed. New York: Macmillan.
[SPPA] Strategic Partnership Program Agroterrorism (SPPA) Initiative. 2007. Available from: http://www.cfsan.fda.gov/~dms/agroterr.html. Accessed Jun 12, 2007.
Wein LM, Liu Y. 2005. Analyzing a bioterror attack on the food supply: the case of botulnium toxin in milk. PNAS 102:9984–9.
Wells WF. 1955. Airborne contagion and hygiene. Cambridge, Mass.: Harvard Univ. Press.

Appendix

This appendix contains the derivation of Eq. 1 to 10 in the main text and briefly discusses the challenges involved in generalizing this approach to microbial agents that undergo growth. All primitive model parameters are described in Table 1 of the main text.

The mean amount of contaminated food. We begin by computing the total amount of contaminated food, H. Let h_i be the amount of contaminated food in stage $i = 1 \ldots, N$, and let m_i denote the number of the n_i containers that are contaminated. By definition, we have $H = h_N$, and this quantity is computed by a simple iterative process. We assume that the deliberate release of agent occurs in one container in stage 1, so that $m_1 = 1$. If stage 1 is a batch process then $h_1 = s_1$. If stage 1 is continuous flow, then the amount of contaminated food at stage 1 depends on when during the cleaning cycle the release occurs (Wein and Liu 2005). By Little's formula (Gross and Harris 1985), h_1 in a continuous flow system is the arrival rate of food at a single container at stage 1, which is $\frac{\lambda_1}{n_1}$, times the length of time between the release time and the next cleaning time. For simplicity, we assume a random release time and ignore the filling and draining intervals at the beginning and end of the cleaning cycle, so that this length of time is uniformly distributed between 0 and w_1. Hence, the expected amount of contaminated food at stage 1 is

$$h_1 = \begin{cases} s_1 & \text{if stage 1 is batch} \\ \dfrac{\lambda_1 w_1}{2n_1} & \text{if stage 1 is continuous flow} \end{cases} \quad \text{(A1)}$$

Regardless of whether stages $2, \ldots, N$ are batch or continuous flow, we can write down simple relations between stages i and $i+1$ in terms of o_i, which is the type of operational flow between stage i and $i+1$ (Figure 1). If stage i is dedicated ($o_i = d$) then the only change is via the volume transformation factor, and $m_{i+1} = m_i$ and $h_{i+1} = \theta_i h_i$. If stage i is forked ($o_i = f$) then no previously uncontaminated food becomes contaminated even though the number of contaminated containers increases, and we have $m_{i+1} = \frac{n_{i+1} m_i}{n_i}$ and $h_{i+1} = \theta_i h_i$. The least straightforward case is assembly ($o_i = a$) because if there were upstream forks then the amount of previously uncontaminated food that becomes contaminated depends on the extent to which the upstream forks correspond to the downstream reassemblies. For example, if at stage 1 a contaminated container forks into streams a and b and an uncontaminated container forks into c and d, and these 4 streams are reassembled into 2 containers at stage 3, then there is no new contaminated food if a and b are assembled together and c and d are assembled together, whereas there is new contaminated food if a is assembled with either c or d. Because upstream forks are typically associated with downstream assemblies in most processing facilities (that is, a and b would be reassembled in our example), we conservatively assume that the reassemblies are performed to minimize the amount of new contaminated food. Under this assumption, we have $m_{i+1} = \lceil \frac{n_{i+1} m_i}{n_i} \rceil$, where $\lceil x \rceil$ is the smallest integer that is greater than or equal to x, and $h_{i+1} = \frac{m_{i+1} n_i}{m_i n_{i+1}} h_i \theta_i$. Finally, if stage i is full mixing ($o_i = m$) then $m_{i+1} = n_{i+1}$ and $h_{i+1} = n_1 h_1 (\prod_{j=1}^{i} \theta_j)$. Solving these equations iteratively from $i = 1$ to $i = N-1$ with initial conditions $m_1 = 1$ and Eq. A1 yields $H = h_N$. Note that because we are only multiplying h_i by a constant at each step of the iteration, we need not keep track of the fact that h_1 is uniformly distributed and can simply use Eq. A1.

The mean number of casualties. Turning to the derivation of Eq. 5 to 7 in the main text, we let time 0 corresponds to the time when contaminated food first leaves the food processing facility; as described in the next paragraph, not all of the contaminated food exits the facility at the same time. We assume that each of the nH potential food consumers has an associated independent and identically distributed version of each of 4 independent random variables.

The 1st random variable is $X^{(1)}$, which is the time interval from when the 1st contaminated food leaves the facility until the time when the contaminated food consumed by a random consumer leaves the facility. We assume that contaminated food leaves the facility at rate $\frac{m_N \lambda_N}{n_N}$, which is the total output rate of food times the fraction of food at stage N that is contaminated. Because the amount of contaminated food is H, if this food exited the facility at a constant rate, then the time that a random potential consumer's contaminated food would leave the facility would be uniformly distributed between 0 and $\frac{n_N H}{m_N \lambda_N}$.

The 2nd random variable is $X^{(2)}$, which is referred to as the speed of the distribution channel. It is the interval from when the contaminated food of a random consumer leaves the facility until it is purchased. We assume $X^{(2)}$ is the sum of a constant δ and a lognormal random variable with median e^{μ_d} and dispersal factor e^{σ_d}.

The 3rd random variable is a uniform [0, 1] random variable u, which leads to the infectious dose $Y = \mathrm{ID}_{50}\, 10^{\Phi^{-1}(u)/\beta}$, where ID_{50} is the median infectious dose, β is the probit slope, and $\Phi(\cdot)$ is the standard normal cdf. Finally, the random variable Z is the incubation period, which is lognormally distributed with median e^{μ_s} and dispersal factor e^{σ_s} (that is, lnZ is a normal random variable with mean μ_s and standard deviation σ_s).

Because αQ amount of active agent is potentially consumed by nH people, in the absence of detection a random potential consumer (indexed by $j = 1, \ldots, nH$) begins consuming food at the time of purchase, denoted by X_j (which equals $X_j^{(1)} + X_j^{(2)}$), and consumes active agent at rate $\frac{\alpha Qc}{nH}$ for c^{-1} time units. Hence, at time $X_j + c^{-1}$ this person would have consumed his entire allotment of active agent, $\frac{\alpha Q}{nH}$. This person would become infected (or poisoned, depending on the type of agent) at time $X_j + \frac{nHY_j}{\alpha Qc}$ if $Y_j \le \frac{\alpha Q}{nH}$; otherwise, he does not become infected. Therefore, in the absence of detection, the number of casualties I is a binomial random variable with parameters nH and $\Phi(\beta \log_{10}(\frac{\alpha Q/nH}{\mathrm{ID}_{50}}))$, and the mean number of casualties is $nH\Phi(\beta \log_{10}(\frac{\alpha Q/nH}{\mathrm{ID}_{50}}))$. We assume that if this quantity is less than k, which is the number of symptomatics needed to identify the attack, then the attack goes undetected and this is the mean number of casualties.

To compute the mean number of casualties if $nH\Phi(\beta\log_{10}(\frac{\alpha Q/nH}{\mathrm{ID}_{50}})) \geq k$, we define the random variable

$$V_j = \begin{cases} \dfrac{nHY_j}{\alpha Qc} & \text{if } Y_j \leq \dfrac{\alpha Q}{nH} \\ \infty & \text{if } Y_j > \dfrac{\alpha Q}{nH} \end{cases} \tag{A2}$$

which represents the amount of consumption time until consumer $j = 1, \cdots, nH$ gets infected. We also let $W_j = X_j + V_j + Z_j$ be the (possibly infinite) time at which potential consumer j develops symptoms, and let m be a binomial random variable with parameters nH and $\Phi(\beta\log_{10}(\frac{\alpha Q/nH}{\mathrm{ID}_{50}}))$, so that m is the number of consumers who will be infected if there is no detection. If we let $W_{(1)} \leq \cdots \leq W_{(m)}$ be the corresponding order statistics, then the time that consumption is halted, which occurs Δ time units after the kth symptomatic patient, is $T = W_{(k)} + \Delta$.

Because the exact calculation of T is burdensome, we derive a tractable but accurate approximation via 2 steps. In the 1st step, we make the simplifying assumption that W is a lognormal random variable with median e^{μ_w} and dispersal factor e^{σ_w}. Because W is the sum of 4 independent variables, we obtain μ_w and σ_w by equating the mean and variance of this lognormal random variable with the mean and variance of $X^{(1)} + X^{(2)} + V + Z$.

The random variable $X^{(1)}$ has mean $\frac{n_N H}{2m_N\lambda_N}$ and variance $\frac{n_N^2 H^2}{12m_N^2\lambda_N^2}$. For $X^{(2)}$, the mean and variance for the nondeterministic part $X^{(2)} - \delta$, which is lognormal, are $e^{\mu_d + \frac{\sigma_d^2}{2}}$ and $(e^{\sigma_d^2} - 1)e^{2\mu_d + \sigma_d^2}$. Similarly, Z has mean $e^{\mu_s + \frac{\sigma_s^2}{2}}$ and variance $(e^{\sigma_s^2} - 1)e^{2\mu_s + \sigma_s^2}$.

For V, recalling that $Y = \mathrm{ID}_{50}\, 10^{\Phi^{-1}(u)/\beta}$, we define the conditional mean $\mu_i = E(Y \mid Y \leq \frac{\alpha Q}{nH})$ and conditional variance $\sigma_i^2 = \mathrm{Var}(Y \mid Y \leq \frac{\alpha Q}{nH})$. This approach, which ignores the fact that some people do not get infected, should be accurate except perhaps in the narrow range where the mean number of casualties is only slightly larger than k. Then the mean and variance for V are $\frac{nH\mu_i}{\alpha Qc}$ and $(\frac{nH\sigma_i}{\alpha Qc})^2$.

Therefore, we obtain μ_w and σ_w by solving the equations

$$e^{\mu_w + \frac{\sigma_w^2}{2}} = \frac{n_N H}{2m_N\lambda_N} + \delta + e^{\mu_d + \frac{\sigma_d^2}{2}} + \frac{nH\mu_i}{\alpha Qc} + e^{\mu_s + \frac{\sigma_s^2}{2}}$$

$$\left(e^{\sigma_w^2} - 1\right)e^{2\mu_w + \sigma_w^2} = \frac{n_N^2 H^2}{12m_N^2\lambda_N^2} + \left(e^{\sigma_d^2} - 1\right)e^{2\mu_d + \sigma_d^2} + \left(\frac{nH\sigma_i}{\alpha Qc}\right)^2 + \left(e^{\sigma_s^2} - 1\right)e^{2\mu_s + \sigma_s^2}$$

In the 2nd step, we use Theorem 5.8 of Balkema and De Haan (1978), which is a limit theorem for $logW_{(k)}$ as $min\{k, m-k\} \to \infty$. Applying this theorem in our setting, we find that $logW_{(k)}$ converges to a normal random variable with mean

$$\mu = \mu_w + \Phi^{-1}\left(\frac{k}{nH}\right)\sigma_w \tag{A3}$$

and standard deviation

$$\sigma = \frac{2\pi}{nH}\sigma_w^2 \frac{k}{nH}\left(1 - \frac{k}{nH}\right)\exp\left(\left(\Phi^{-1}\left(\frac{k}{nH}\right)\right)^2\right) \tag{A4}$$

To prevent the need for multiple integration in our final formula for the case in which the attack is detected, we hereafter assume that consumption is halted at the deterministic time $\tau = e^{\mu} + \Delta$. Our numerical computations suggest that the omission of A4 is inconsequential because $\sigma \ll \mu$.

With τ in hand, we derive the amount of active agent consumed by consumer j as follows. Suppose for consumer j, x_1 is the time for the food to leave the facility, and x_2 is the nondeterministic part in the distribution channel. Then consumer j starts consumption at time $x_1 + \delta + x_2$, consumes the amount

$$C_j = \frac{\alpha Qc\, \max\left(0, \min\left(c^{-1}, \tau - \left(x_1 + \delta + x_2\right)\right)\right)}{nH} \tag{A5}$$

and gets infected with probability $\Phi(\beta\log_{10}(\frac{C_j}{\mathrm{ID}_{50}}))$. Using the probability density functions for the uniform x_1 and lognormal x_2, we get

$$E(I) = nH\int_0^{\infty}\int_0^{\frac{n_N H}{m_N\lambda_N}} \Phi\left(\beta\log_{10}\left(\frac{C_j}{\mathrm{ID}_{50}}\right)\right) \times \frac{m_N\lambda_N}{n_N H}\frac{1}{x_2\sigma\sqrt{2\pi}}e^{-\frac{(\ln x_2 - \mu_d)^2}{2\sigma_d^2}}\, dx_1\, dx_2 \tag{A6}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution.

In summary, the mean number of casualties is

$$E[I] = nH\Phi\left(\beta\log_{10}\left(\frac{\frac{\alpha Q}{nH}}{\mathrm{ID}_{50}}\right)\right) \quad \text{if } nH\Phi\left(\beta\log_{10}\left(\frac{\frac{\alpha Q}{nH}}{\mathrm{ID}_{50}}\right)\right) < k \tag{A7}$$

and is expressed by Eq. A6 otherwise, where $\tau = e^{\mu} + \Delta$ and μ is given by A3.

Microbial growth. Finally, we briefly outline how to generalize our model to a microbial agent that undergoes growth via a standard function such as the Gompertz function (for example, Table 1 in Buchanan [1991]) or the Baranyi–Roberts function (Baranyi and Roberts 1994). For simplicity, we assume an amount Q of agent is introduced (deliberately or accidentally) into 1 container in stage 1. If we assume that the amount of time that food spends at each of the N stages is a constant, then the concentration of the agent at each point in time (starting from its initial concentration in stage 1 of $\frac{Q}{s_1}$) can be computed by assuming it follows the specified growth function during each processing stage in addition to instantaneous dilutions (that is, multiplying the concentration by $\frac{h_{i-1}}{h_i}$) when the food is transferred from stage $i-1$ to stage i. The remainder of the upstream analysis remains the same, and we denote the concentration and volume of contaminated food at the end of stage N by K and H, respectively.

Moving to the downstream portion of the supply chain, if we let time 0 be the time when contaminated food first leaves the food processing facility and let $K(t)$ be the concentration of contaminated food at time t, then $K(t)$ continues to undergo growth (for example, Gompertz or Baranyi–Roberts) with the initial condition $K(0) = K$. As before, each of the nH potential consumers has 4 i.i.d. random variables: the time in the facility $X^{(1)}$, the distribution speed $X^{(2)}$, the infection threshold Y_j, and the incubation period Z_j. Also, as before, poisoned consumers exhibit symptoms at time $\tilde{W}_j = X_j^{(1)} + X_j^{(2)} + \tilde{V}_j + Z_j$, where $\tilde{V}_j$ is the time needed to consume the required amount of food to cause infection if food is consumed at rate c forever. The key challenge is to compute the mean and variance of $\tilde{W}_j$, after which the analysis is identical to the steps taken to derive Eq. A3 to A7. However, the derivation of the mean and variance of $\tilde{W}_j$ is more difficult than before because $\tilde{V}_j$ depends not only on Y_j, but also on $X_j = X_j^{(1)} + X_j^{(2)}$.

The value of $\tilde{V}_j$ is derived by numerically solving

$$\int_{X_j}^{X_j+\tilde{V}_j} K(t)c\,dt = Y_j \tag{A8}$$

Then the generalization of Eq. A2 is

$$V_j = \begin{cases} \tilde{V}_j & \text{if } \tilde{V}_j \le \dfrac{1}{nc} \\ \infty & \text{if } \tilde{V}_j > \dfrac{1}{nc} \end{cases} \tag{A9}$$

If we drop the subscript j and denote the solution to Eq. A8 by $\tilde{V}(x, y)$, then the mean and variance of $X + \tilde{V}$ are, respectively,

$$\iint (x + \tilde{V}(x, y))\, f(x)\, p(y)\; dx\, dy \tag{A10}$$

$$\iint (x + \tilde{V}(x, y))^2\, f(x)\, p(y)\; dx\, dy - \left(\iint (x + \tilde{V}(x, y))\, f(x)\, p(y)\; dx\, dy\right)^2 \tag{A11}$$

Computing Eq. A10 and A11 numerically and adding the mean $e^{\mu_s + \frac{\sigma_s^2}{2}}$ and variance $(e^{\sigma_s^2} - 1)e^{2\mu_s + \sigma_s^2}\mu_s$ of Z_j to Eq. A10 and A11, respectively, yields the mean and variance of $\tilde{W}_j$.