Supplementary information

A The single hit model for infection

A.1 Exposure

Simplest we can assume for exposure is Poisson sample, from a suspension of strength C, sample volume V

$$\text{Prob}(n|C,V) = \frac{(CV)^n}{n!}\text{e}^{-CV} \tag{A.1}$$

Where CV is the (expected) dose, the probability of exposure

$$\text{Prob}(n \geq 1|C,V) = 1 - \text{Prob}(n = 0|C,V) = 1 - \text{e}^{-CV} \tag{A.2}$$

A.2 Infection: fixed p_m

Suppose we have a host who has ingested n pathogens, and all pathogens have equal survival probabilities p_m, then the probability that k pathogens survive is

$$\text{Prob}(k|n,p_m) = \binom{n}{k} p_m^k (1-p_m)^{n-k} \tag{A.3}$$

if survival is independent.

Infection corresponds to survival of at least 1 pathogen (a 'single hit') with probability

$$\text{Prob}(k \geq 1|n,p_m) = 1 - \text{Prob}(k = 0|n,p_m) = 1 - (1-p_m)^n \tag{A.4}$$

The marginal probability of infection therefore is

$$\text{Prob}(k \geq 1|C,V,p_m) = \sum_{n=1}^{\infty} \frac{\text{e}^{-CV}(CV)^n}{n!}\left[1-(1-p_m)^n\right] \tag{A.5}$$

which can be simplified, by first taking the sum from $n = 0$

$$P_{\text{inf}}(C,V,p_m) = 1 - \sum_{n=0}^{\infty} \frac{\text{e}^{-CV}(CV)^n}{n!}(1-p_m)^n \tag{A.6}$$

and noting that

$$\sum_{n=0}^{\infty} \frac{\text{e}^{-CV(1-p_m)}\left[CV(1-p_m)\right]^n}{n!} = 1 \tag{A.7}$$

so that

$$P_{\text{inf}}(C, V, p_m) = 1 - \mathrm{e}^{-CV}\mathrm{e}^{CV(1-p_m)} = 1 - \mathrm{e}^{-p_m CV} \tag{A.8}$$

A.3 Infection: heterogeneous p_m

In case of heterogeneity in p_m, described by a Beta pdf

$$f(p_m|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p_m^{\alpha-1}(1-p_m)^{\beta-1} \tag{A.9}$$

the marginal dose response relation for infection becomes

$$P_{\text{inf}}(C, V|\alpha, \beta) = \int_{p_m=0}^{\infty} f(p_m|\alpha, \beta)\left(1 - \mathrm{e}^{-p_m CV}\right) \tag{A.10}$$

which can be written as a (Kummer) confluent hypergeometric function

$$P_{\text{inf}}(C, V|\alpha, \beta) = {}_1F_1(\alpha, \alpha+\beta; -CV) \tag{A.11}$$

Furumoto and Mickey [23] have shown how this relation can be simplified into

$$P_{\text{inf}}(C, V|\alpha, \beta) = 1 - \left(1 + \frac{CV}{\beta}\right)^{-\alpha} \quad (\beta \gg 1; \alpha \ll \beta) \tag{A.12}$$

B Heterogeneity in the dose

B.1 Exposure, dose variable

In outbreak situations the dose often is inappropriately characterized by a simple Poisson model. Instead, we may use a Poisson-Gamma mixture to model extra–Poisson variation. The observed number is again a Poisson sample

$$\text{Prob}(n|C, V) = \frac{(C \cdot V)^n}{n!}\mathrm{e}^{-C \cdot V} \tag{A.13}$$

Where $C \cdot V$ is the (expected) dose. The concentration C now is assumed to have a Gamma density

$$g(C|\rho, \lambda) = \frac{\lambda^{-\rho}}{\Gamma(\rho)} C^{\rho-1}\mathrm{e}^{-C/\lambda} \tag{A.14}$$

with shape parameter ρ and scale parameter λ. The marginal distribution of the counts then is negative binomial

$$\text{Prob}(n|\rho,\lambda,V) = \frac{\Gamma(n+\rho)}{n!\Gamma(\rho)}\left(\frac{1}{1+\lambda V}\right)^{\rho}\left(1-\frac{1}{1+\lambda V}\right)^{n} \quad \text{(A.15)}$$

And the probability of exposure is

$$\text{Prob}(n \geq 1|\rho,\lambda,V) = 1-\text{Prob}(n=0|\rho,\lambda,V) = 1-(1+\lambda V)^{-\rho} \quad \text{(A.16)}$$

which may be written as

$$\text{Prob}(n \geq 1|\rho,\tilde{c},V) = 1-\left(1+\frac{\tilde{c}V}{\rho}\right)^{-\rho} \quad \text{(A.17)}$$

where $\tilde{c} = \lambda\rho$ is the mean concentration.

B.2 Infection, dose variable, fixed p_m

The marginal probability of infection can be found, as above

$$\text{Prob}(k \geq 1|\rho,u,p_m) = \sum_{n=1}^{\infty}\frac{\Gamma(n+\rho)}{n!\Gamma(\rho)}u^{\rho}(1-u)^{n}\left[1-(1-p_m)^{n}\right] \quad \text{(A.18)}$$

substituting $u = 1/(1+\lambda V)$.

This can be simplified by first taking the sum from $n = 0$

$$P_{\text{inf}}(\rho,u,p_m) = 1-\sum_{n=0}^{\infty}\frac{\Gamma(n+\rho)}{n!\Gamma(\rho)}u^{\rho}(1-u)^{n}(1-p_m)^{n} \quad \text{(A.19)}$$

If we note that

$$\sum_{n=0}^{\infty}\frac{\Gamma(n+\rho)}{n!\Gamma(\rho)}\left[1-(1-u)(1-p_m)\right]^{\rho}\left[(1-u)(1-p_m)\right]^{n} = 1 \quad \text{(A.20)}$$

then

$$P_{\text{inf}}(\rho,u,p_m) = 1-\left(\frac{1-(1-u)(1-p_m)}{u}\right)^{-\rho} \quad \text{(A.21)}$$

or

$$P_{\text{inf}}(\rho,\lambda,V,p_m) = 1-(1+\lambda V p_m)^{-\rho} \quad \text{(A.22)}$$

$$P_{\text{inf}}(\rho,\tilde{c},V,p_m) = 1-\left(1+\frac{\tilde{c}V}{\rho}p_m\right)^{-\rho} \quad \text{(A.23)}$$

B.3 Infection: heterogeneous p_m

In case of heterogeneity in p_m, described by a Beta pdf the marginal dose response relation for infection becomes

$$P_{\text{inf}}(\rho, \tilde{c}, V|\alpha, \beta) = \int_{p_m=0}^{\infty} f(p_m|\alpha, \beta) \left(1 - \left(1 + \frac{\tilde{c}V}{\rho} p_m\right)^{-\rho}\right) \quad \text{(A.24)}$$

which can be written as another hypergeometric function

$$P_{\text{inf}}(\rho, \tilde{c}, V|\alpha, \beta) = {}_2F_1(\alpha, \rho, \alpha + \beta; -\tilde{c}V/\rho) \quad \text{(A.25)}$$

C Sexual reproduction and infection

Suppose we have a host who has ingested of n pathogens, of whom k females (♀) and $n - k$ males (♂).

Infection can occur if and only if 1 or more ♀ pathogens and 1 or more ♂ pathogens survive. Suppose ♀ and ♂ pathogens have equal survival probabilities p_m, then the probability that 1 or more ♀ pathogens survive is

$$p_{♀} = 1 - (1 - p_m)^k \quad \text{(A.26)}$$

and the probability that 1 or more ♂ pathogens survive

$$p_{♂} = 1 - (1 - p_m)^{n-k} \quad \text{(A.27)}$$

Suppose ♀ and ♂ pathogens are present in proportions r and $1 - r$ (r is the sex ratio: the fraction ♀). Then the numbers of ♀ and ♂ pathogens are binomial

$$\text{Prob}(k♀, n - k♂|r) = \binom{n}{k} r^k (1 - r)^{n-k} \quad \text{(A.28)}$$

and the probability of infection is

$$P_{\text{inf}}(n|r) = \sum_{k=0}^{n} \binom{n}{k} r^k (1 - r)^{n-k} \left[1 - (1 - p_m)^k\right] \left[1 - (1 - p_m)^{n-k}\right] \quad \text{(A.29)}$$

which can be shown to equal

$$P_{\text{inf}}(n|r) = 1 + (1 - p_m)^n - [1 - p_m(1 - r)]^n - (1 - p_m r)^n \quad \text{(A.30)}$$

C.1 Exposure

Simplest we can assume for exposure is Poisson sample, from a suspension of strength C, sample volume V. The exposure dose response relation (de-

scribing the probability of having ingested at least 1 ♀and 1♂organism) can be written as a linear combination of three terms

$$\text{Prob}(♀ \geq 1, ♂ \geq 1|r, C, V) = 1 - \mathrm{e}^{-(1-r)CV} - \mathrm{e}^{-rCV} \tag{A.31}$$

simply by taking the terms of equation (A.30) for $p_m = 1$.

For heterogeneous exposure we can again assume a Poisson–Gamma mixture, leading to an exposure dose response

$$\text{Prob}(♀ \geq 1, ♂ \geq 1|r, \rho, \tilde{c}, V) = 1 - \left(1 + \frac{\tilde{c}V}{\rho}(1-r)\right)^{-\rho} - \left(1 + \frac{\tilde{c}V}{\rho}r\right)^{-\rho} \tag{A.32}$$

C.2 Infection, fixed p_m

For Poisson exposure and fixed "hit" probability p_m we get

$$P_{\text{inf}}(C \cdot V|p_m, r) = 1 + \mathrm{e}^{-C \cdot V p_m} - \mathrm{e}^{-C \cdot V p_m (1-r)} - \mathrm{e}^{-C \cdot V p_m r} \tag{A.33}$$

analogous to the exponential dose response relation for asexually reproducing pathogens.

For Poisson–Gamma exposure the relation is

$$P_{\text{inf}}(\rho, \tilde{c}, V|p_m, r) = 1 + \left(1 + \frac{\tilde{c}V}{\rho}p_m\right)^{-\rho} - \left(1 + \frac{\tilde{c}V}{\rho}p_m(1-r)\right)^{-\rho} - \left(1 + \frac{\tilde{c}V}{\rho}p_m r\right)^{-\rho} \tag{A.34}$$

C.3 Infection, variable p_m

The model for heterogeneous p_m can again be written as a linear combination of hypergeometric relations (see equation (A.11)).

$$P_{\text{inf}}(C \cdot V|\alpha, \beta, r) = 1 + {}_1F_1\left[\alpha, \alpha+\beta; -C \cdot V\right] - {}_1F_1\left[\alpha, \alpha+\beta; -C \cdot V(1-r)\right] - {}_1F_1\left[\alpha, \alpha+\beta; -C \cdot Vr\right] \tag{A.35}$$

In case the dose also has extra–Poisson variation, the resulting dose response relation is a combination of the functions in equation (A.25)

$$P_{\text{inf}}(\rho, \tilde{c}, V|\alpha, \beta, r) = 1 + {}_2F_1(\alpha, \rho, \alpha+\beta; -\tilde{c}V/\rho) - {}_2F_1(\alpha, \rho, \alpha+\beta; -\tilde{c}V(1-r)/\rho) - {}_2F_1(\alpha, \rho, \alpha+\beta; -\tilde{c}Vr/\rho) \tag{A.36}$$

D Hierarchical dose response model

The likelihood is binomial: for each incident K out of N subjects exposed to a dose $D = g(\rho, \tilde{c}, V)$ have been observed to be affected.

Given the hit theory dose response function a single observed attack rate may allow prediction of the dose response relation [2]. We want to incorporate multiple attack rates at various doses. However, such an approach inevitably involves an additional level of biological variation. While a different human population similar in age and health status might have similar susceptibility, a different isolate of the pathogen is likely to have completely different infectivity, if only because of a different history (different food vehicle, different previous host, ...). Therefore, analysis of data from different outbreaks requires a hierarchical model (Figure A1).

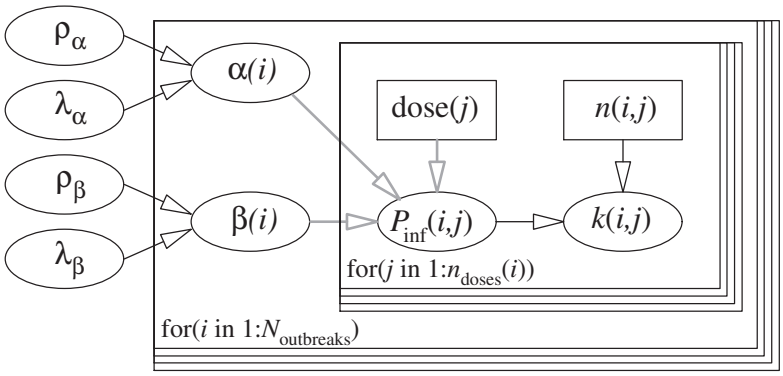

Figure A1: Two-level model for dose response assessment of several outbreaks, each with their separate pathogen isolates and possibly susceptibility distributions ($N_{\text{outbreaks}}$ = number of outbreaks; $n_{\text{doses}}(j)$ = 1 for all outbreaks except the first [5], where $n_{\text{doses}}(1) = 2$).

If there are j observations in group i and the dose response model

$$f(d|\boldsymbol{\theta}) \tag{A.37}$$

with parameter vector $\boldsymbol{\theta}$ the contribution of group i to the likelihood is

$$\ell_i(\boldsymbol{\theta}) = \prod_j [f(d_{i,j}|\boldsymbol{\theta})]^{k_{i,j}} [1 - f(d_{i,j}|\boldsymbol{\theta})]^{n_{i,j}-k_{i,j}} \tag{A.38}$$

all observations in group i share the same parameter set $\boldsymbol{\theta}$

When the joint distribution of $\boldsymbol{\theta}$ over all groups is

$$h(\boldsymbol{\theta}|\boldsymbol{\Xi}) \tag{A.39}$$

with hyperparameter vector $\boldsymbol{\Xi}$, the marginal likelihood can be written

$$\begin{aligned}\ell_i(\boldsymbol{\Xi}) &= \int_{\boldsymbol{\theta}} \ell_i(\boldsymbol{\theta}) h(\boldsymbol{\theta}|\boldsymbol{\Xi}) \mathrm{d}\boldsymbol{\theta} \\ &= \int_{\boldsymbol{\theta}} \left[f(d_{i,j}|\boldsymbol{\theta})\right]^{k_{i,j}} \left[1 - f(d_{i,j}|\boldsymbol{\theta})\right]^{n_{i,j}-k_{i,j}} h(\boldsymbol{\theta}|\boldsymbol{\Xi}) \mathrm{d}\boldsymbol{\theta} \quad \text{(A.40)}\end{aligned}$$

and the hierarchical likelihood, to be evaluated, is

$$\mathrm{L}(\boldsymbol{\Xi}) = \prod_i \ell_i(\boldsymbol{\Xi}) \quad \text{(A.41)}$$

The dose is characterized by the expected concentration of pathogens, and their variation, characterized by the Gamma shape parameter ρ. These two parameters are estimated separately using whatever information was available in the outbreak reports, usually quantiles characterizing location and spread of intake of contaminated unheated (or inadequately heated) meat.

Infectivity parameters are transformed as in [2]: since we have only one data point per outbreak, the parameters (α, β) of the Beta Poisson model are highly correlated: parameter estimation is improved by transformation to

$$\begin{aligned} u &= \alpha/(\alpha+\beta) \\ v &= {}^{10}\log(\alpha+\beta) \end{aligned} \quad \text{(A.42)}$$

so that we are estimating the mean value (u) of the Beta distribution for p_m and a quantity that is inversely related to its variance (for very large positive values of v the variance tends to zero). Further u is logit–transformed and v is log–transformed

$$\begin{aligned} w &= \log[u/(1-u)] \\ z &= \log(v) \end{aligned} \quad \text{(A.43)}$$

We use normal priors for w and z (mean ρ, standard deviation λ). Uncorrelated non–informative normal (-8,8) hyperpriors were taken for the means of w and z (rho), gamma (0.001,1000) priors were taken for the standard deviations of w and z (λ).

Posterior parameter samples have been obtained using the Metropolis-Hastings algorithm, implemented in Mathematica [2].

E Additional outbreak information

City/village name	year	nr. cases	parasite strain	asympt. cases	type of meat	cons. g/pers.	conc. larv./g	cons. abroad	meat import	animal import	local	reference
Perpignan	2001	1	NA	NA	pork	NA	NA	yes				CNR *Trichinella*
Paris	2001	1	NA	NA	NA	NA	NA	yes				CNR *Trichinella*
Quillan (Aude)	2002	4	NA	NA	Wild Boar	NA	NA				yes	CNR *Trichinella*
Villeneuve d'Entraunes (Alpes Maritimes)	2003	6	*T. britovi*	1	Wild Boar	150	3				yes	[7]
Rouen (Seine Maritime)	2004	1	*T. nativa*	0	Black Bear	300	250	yes				CNR *Trichinella*
Martigues (Bouches du Rhône)	2004	1	*T. britovi*	0	Jackal	NA	NA	yes				[24]
	2004	1	NA	NA	NA	NA	NA	yes				CNR *Trichinella*
Rouen (Seine maritime)	2005	3	NA	6	pork	NA	NA	yes				CNR *Trichinella*
Orléans (Loiret), Narbonne (Aude)	2005	17	*T. nativa*	8	Black Bear	150	300	yes	yes			[25]
Ore (31)	2006	2	NA	NA	Wild Boar	NA	NA				yes	CNR *Trichinella*
Nans les Pins (Var)	2006	3	*T. spiralis*	NA	Wild Boar	200	40					CNR *Trichinella*
Collobriëres (Var)	2006	4	*T. britovi*	6	Wild Boar	150	5–10				yes	CNR *Trichinella*

Table A1: Data from France provided by J. Dupouy–Camet (May 2007).