Annex C

Initial Contamination of *Salmonella* Enteritidis in Shell Eggs

INTRODUCTION

This annex discusses the methods used to estimate the levels of *Salmonella* Enteritidis (SE) in contaminated eggs at lay. These estimates are based on a conceptual model of transovarian infection (vertical transmission) of SE from hen to egg. SE contamination of the hen reproductive system can lead to contamination of the internal contents of shell eggs with different levels of SE. It is generally believed that these initial levels are low, yet high initial levels of contamination can occur. SE contaminating the hen ovary and/or oviduct can be deposited at different locations within the egg, i.e. yolk and/or albumen. Depending where SE are deposited, their ability to grow and therefore their likelihood of detection will vary. This annex estimates the initial levels of SE in egg albumen and egg yolk.

To estimate the levels of SE in eggs at lay, studies enumerating the levels of SE in eggs from hens experimentally inoculated with SE were evaluated. A primary concern in using these datawasthe effect of the contamination mechanism used to infect the egg (migration of direct deposition), the age of the egg at collection, the time-temperature history of an egg, and the location of contamination are not known. These variables make it difficult to determine whether

the enumerated levels of SE are due to initial contamination levels or due to growth that might have taken place between the time of contamination and the time that the egg was analyzed. Consequently, assumptions concerning how much growth might have occurred if enumeration was not immediate are included. These assumptions define the model and enable estimates to be made of values for the parameters for this model.

This annex is divided into two main sections: 1) distribution of the levels of SE in albumen (*Ea*) contaminated eggs; and 2) distribution of levels of SE in the yolk (*Ev*) and vitelline membrane[a] (*Ev*). These sections estimate the distribution of initial levels of *Ea* and *Ey* contaminated eggs accompanied by the standard deviation and uncertainty surround these distributions. The distribution of SE in *Es* eggs and *Salmonella* spp. in *Ep* eggs are assumed to be the same as that in *Ea* eggs. This assumption wasa "state of knowledge" assumption made in the absence of data.

DISTRIBUTION OF LEVELS OF SE IN *EA* CONTAMINATED EGGS AT TIME OF LAY

Estimating the Distribution of the Number of Cells in Albumen

SE can be deposited within the albumen of an egg by vertical transmission. SE can grow to substantial levels in egg albumen, thereby posing a risk to the consumer. Contamination of the albumen (*Ea*) is thought to be the primary site of SE infection within eggs, and SE from *Ea* infections can migrate to other egg compartments. Therefore, determining the initial levels of SE within the albumen is important to determine risk.

To estimate the initial levels of SE in egg albumen, we used raw data provided by Gast[1] and Cogan.[2] These data were from studies originally published by Gast and Beard[3] and Cogan et al.,[4] respectively.

The study of Gast and Beard[3] was used for determining initial levels of SE in albumen. This study enumerated SE from contaminated eggs from experimentally inoculated hens. These eggs were collected daily between 4 and 14 days post-inoculation. Three handling procedures were used for eggs produced by these treated hens: Group 1) contents of the eggs were analyzed for SE the same day eggs were collected; Group 2) eggs were stored at 7.2 C for 7 days before analyzing; Group 3) eggs were stored at 25 C for 7 days before analyzing.

Following these handling procedures, albumen samples were taken from eggs and frozen. Later, samples were thawed and analyzed for the presence of SE. The raw data from this study was obtained through personal communication[1] and are presented in Tables C1 and C2. An analysis of these data is presented in Table C3.

[a] Vitelline membrane is the surrounding yolk membrane. This membrane can be contaminated with SE without yolk contamination.

TABLE C1 GROUP 1[a] AND 2[b] COMBINED RAW DATA.[1]

	Measured SE cfu/mL						
	2	**3**	**4**	**6**	**7**	**8**	**9**
Frequency of Observations	2	1	2	1	1	1	1

[a]Group 1 sampled on day of lay
[b]Group 2 stored at 7.2 °C and sampled seven days after lay

Table C2 Group 3[a] raw data.[1]

	Measured SE cfu/mL												
	1	**2**	**3**	**4**	**5**	**7**	**8**	**9**	**12**	**35**	**38**	**85**	**106**
Frequency of observations	4	2	2	4	1	1	2	1	1	1	1	1	1

[a]Group 3 stored at 25 °C and sampled 7 days after lay

As the age of the eggs was unknown, the enumerated values for Groups 1 and 2 (Table C1) might represent initial contamination levels or subsequent SE growth from unknown levels. Therefore, to estimate the initial levels of SE from this study, relative growth that might have occurred during the time prior to analysis of the eggs was considered. From that, the possible amount of growth within the albumen was determined. The Group 3 data (Table C2) will not be explicitly used as there was a strong possibility that growth occurred in these eggs due to holding at 25 °C for 7 days. These data are used implicitly to determine the percentage of SE-positive eggs (see below; Table C3).

Assigning a Distribution to Initial Contamination in Albumen

For estimating the distribution of the number of SE cells, k, in an egg at lay, it was assumed that the distribution can be written as:

$$\begin{aligned} f(0) &= p && \text{if } x = 0 \\ f(k) &= (1-p)\,g(k \mid \psi) && \text{otherwise.} \end{aligned} \qquad \text{(C1)}$$

where $g(k)$ is the probability density function of k when k is positive, depending upon unknown parameters ψ. It was assumed further that the number of SE cells in an egg at the time of measurement, y, is the product of the initial number of SE cells in the egg, k, and factor, $r(k)$, describing the relative growth, so that $E(y|r, k) = r(k)k$, and the distribution of y given r and k was assumed to be distributed as $b(y|r, k)$. The distribution of $m(y)$, for $k > 0$, can be written as:

$$m(y|\psi, r) = \int b(y|r,k)dg(k|\psi) \quad \text{(C2)}$$

where the integral represents the summation when the distribution is assumed discrete. Thus, the likelihood of an observation, z, depends upon r and k,

$$Lik(z|\psi, r) = p\delta(z=0) + (1-p)\int_0^\infty m(y|\psi, r)h(z|y)dy \quad \text{(C3)}$$

where δ (true expression) = 1 and δ (false expression) = 0, and where $h(z|y)$ is the density of the Poisson distribution with an expected value = λy.

Gast and Beard[3] froze egg samples prior to enumeration. The freeze-thaw process killed about 25% of the SE within the sample.[1] Therefore, the recovery of SE was approximately 75%. To enumerate samples, 5 samples of 0.2 mL albumen were analyzed.[3] Thus it is assumed 1 mL of albumen in total was analyzed. It was also reported that the average volume of the eggs was 40 mL. Thus, $\lambda = 0.75/40 = 1/53.33$.

Finally, it was assumed that r is a random variable with cumulative density function (cdf) $V(r)$, which may depend upon x so that the unconditional likelihood of an observation z, given x and unknown parameters is:

$$Lik(z|\psi) = p\delta(z=0) + (1-p)\int_{r=1}^\infty \int_{y=0}^\infty m(y|\psi, r)h(z|y)dydV(r). \quad \text{(C4)}$$

The distribution V represents the variation of relative growth among SE-contaminated eggs.

Estimating the Relative Growth of SE in Albumen

To use the data of Gast and Beard[3] to determine initial levels of SE in albumen, the level of relative SE growth needed to be determined to account for possible growth within a contaminated egg prior to collection. To estimate this, assumptions are made regarding the physiological state of the SE cells and the age of the eggs. Both factors will affect the relative growth of SE within albumen.

To determine a possible amount of growth of SE within albumen, the growth in isolated egg albumen data from Cogan[2] are available. Cogan isolated egg albumen and inoculated 25 mL samples with varying levels of SE and incubated for 20 and 30 °C for 8 days. Levels of SE/mL were then enumerated. (These data are discussed in further detail within Annex E.) For each egg

sample, an exponential growth rate could be estimated from these data when assuming a fixed lag to generation time ratio, *Rat* (see Annex E). However, as indicated by Cogan et al.,[4] the SE cells used in this study were in stationary phase. SE within naturally contaminated eggs would not likely be in stationary phase (see Annex E), and thus the amount of growth that would be expected for stationary phase cells is less than that for naturally contaminating SE cells. There are no data to describe the growth phase of naturally contaminating SE, or to estimate the effect of this growth phase on SE growth in albumen. Therefore, it was assumed that the exponential growth rates would be the same for both initial phases of SE cells.

Eggs were collected in the morning for the Gast and Beard study.[1,3] Thus, it is possible that some of the collected eggs on a given day were up to 24 hours old. Alternatively, some of the collected eggs could have been laid shortly before they were collected. Therefore, it was assumed that the ages of the eggs in the study were uniformly distributed between 0 and 1 day old when sampled.

Through simulation of data from Cogan,[2] selecting a random time between 0 and 1 for each sample, and computing the amount of SE growth for the sample assuming *Rat* = 1, it was computed that $E(r) = 1.5$ and the standard deviation is 0.65 (so that $E(r^2) = 2.7$). Thus, from this model, the amount of SE growth that samples with more than a few SE cells would experience is not large, generally less than a 3-fold increase (0.5 $\log_{10}$).

SE Density and Growth in Albumen

In a population of eggs inoculated within the albumen, some of the eggs will experience SE growth and others will not. Given favorable time and temperature conditions, this could be due to location of the inoculum relative to the yolk, physiological egg variation, pH, and inoculum size. In regard to inoculum size, smaller inoculums of SE are less likely to grow within the albumen as compared to larger sizes. This is not only from stochastic considerations, but also from microbiological ones, such as forming cell aggregates and sensing the presence of surrounding bacteria, behaviors that may influence bacterial density and alter gene expression. Therefore, to determine the initial levels of SE deposited within the albumen, consideration for those percentages of eggs that did not show SE growth is needed.

Cogan[2] inoculated whole eggs within the albumen close to the yolk (*Eac*) with varying levels of SE and held these eggs at 20 or 30 ºC for 8 days. For many inoculated eggs, SE growth was not observed. This suggests that the distribution of relative SE growth among contaminated eggs, V, of the function r, be assumed such that $V(1)(k) = v_0(k) > 0$. With this notation, the likelihood of z can be written as:

$$Lik(z \mid \psi) = p\delta(z=0) + (1-p)\Big[\int_{k=0}^{\infty} h(z \mid k) v_0(k) dg(k \mid \psi) + \int_{r=1}^{\infty}\int_{k=0}^{\infty} h(z \mid kr)(1-v_0(k)) dg(k \mid \psi) dV(r)\Big] \quad \text{(C5)}$$

Equation C5 expresses that the likelihood of SE growth depends upon the number of SE cells, *k*, at time *t* in the initial contamination. For the inoculation levels of 2, 25, 250, and 2500 SE cells per egg, the percentages of samples for which the relative growth was less than 1.25 were 41, 19, 12 and 12%, respectively, nearly independent of the levels *k*, for sufficiently large *k*. Though there is a trend toward growth dependence and concentration, this result was unexpected based on the above discussion. However, these percentages include eggs in which SE appear to have had access to the yolk nutrients due to yolk membrane breakdown (YMB) (see Annex E). To identify the percentages of samples for which the relative growth was less than 1.25 in the albumen only, eggs for which there were indications of YMB were not included. Upon so doing, the percentages of samples for which relative growth were less than 1.25 increased to 50, 25, 17 and 24% for inoculation levels of 2, 25, 250, and 2500 SE cells per egg, respectively. Once again, for the samples with inoculation of 2 SE cells per egg, the percentage of eggs showing no growth is large, as would be expected; however, for sufficiently large *k*, the levels appear independent of inoculum size**.** Further, for samples with a target inoculum of 2 cells per egg, a number of eggs did not contain SE. Therefore, errors in the number of SE cells that were actually inoculated could be large.

Most of the derived distributions of the number of SE cells for a target inoculum of 2 and 25 cells per egg is below 15 cells, so that it could be reasonably assumed that 20% or so of the eggs would not experience SE growth. This number, as a rough approximation, was assumed for all albumen contaminated (*Ea*) eggs, so that, for the estimation of the distribution *g*, it was assumed that v_0 and *V* are not dependent on *k*. With this simplification, from the above equations, the first and second non-central moments of *z* are:

$$E(z) = (1 - p)(\lambda E(k))(v_0 + (1 - v_0)E(r)) \tag{C6}$$

and:

$$E(z^2) = E(z) + (1 - p)\lambda^2(v_0 E(k^2) + (1 - v_0)E(k^2)E(r^2)) \tag{C7}$$

respectively.

To complete this analysis, 1-*p*, the percentage of eggs contaminated with SE in the albumen, was calculated from Gast and Beard.[3] Hens were experimentally inoculated with SE and eggs were collected daily between 4 and 14 days post-inoculation. Three handling procedures were used for eggs produced by these treated hens: Group 1) contents of the eggs were analyzed for SE the same day eggs were collected; Group 2) eggs were stored at 7.2°C for 7 days before analyzing; Group 3) eggs were stored at 25°C for 7 days before analyzing.

Following these handling procedure, albumen samples were taken from eggs and frozen. Upon enumeration, samples were thawed and analyzed for the presence of SE.

TABLE C3 PERCENT SE POSITIVE SAMPLES AND MEAN SE LEVELS (CFU/mL) FOR THREE GROUPS OF EGGS.

	Group 1[a]	Group 2[b]	Group 1 and 2	Group 3[c]
Number of Positive Samples (%)	4/132 (3)	5/134 (3.7)	9/266 (3.4)	22/138 (15.9)
Mean SE CFU/ml	5.5	4.6	5.0	15.6
Geometric Mean	5.1	3.9	4.4	5.6
Median Value	5.5	4	4	4
25th Percentile[d]	3.25	2	2.5	2
High Value	8	9	9	106

[a]Group 1 sampled on day of lay.
[b]Group 2 stored at 7.2 °C and sampled 7 days after lay.
[c]Group 3 stored at 25 °C and sampled 7 days after lay.
[d]Computed as $y = (1-f)x_i + fx_{i+1}$, where x_i is the i^{th} ordered result, and the number of observations can be expressed as: $4(i+f) - 1$.

Table C3 shows a striking difference in the percentages of positive samples between those of the first two groups and that of the third group, i.e. 3.4% vs. 15.9%, respectively. A comparison between the first two groups and group 3 suggests that for SE deposited in many of the eggs from groups 1 and 2; initial levels of SE were below the detection limit of the methodology used in the study. In addition, a comparison between group 2 stored at 7.2 °C and 3 stored at 25 °C suggests SE deposited in group 3 grew to levels above the detection limit and were therefore recovered more frequently than SE in group 2 eggs held below optimal growth temperatures.

For group 3, approximately 16% (22/138) of the egg samples were positive for SE. The actual percentage of SE positive eggs in this experiment, however, could be greater than reported due to false negatives. From the data of Cogan,[2] 22% of the egg samples (after an 8 day storage at 20 or 30 °C) for which it was determined that there was no YMB event showed less than a 25% increase in the number of SE cells (excluding the results from the lowest inoculation group of 2 cells). Because eggs from the study by Cogan[2] were assumed to represent *Eac* eggs, it is possible that there would be a larger percentage of no growth for *Eaf* eggs. For example, if 22% of the *Eac* eggs and 30% of the *Eaf* eggs did not experience growth, and 70% of the *Ea* eggs are *Eaf*, then the percentage of no growth contaminated egg would be about 28%. Hence for group 3 it is possible that 22% (16%/(1-0.28)) of the eggs were positive for SE. That is, it is possible, that 20% or more of the eggs were positive in groups 1 and 2, of which, only 3-4% of them were detected positive, for approximately an 80 to 85% false negative rate.

Equations C6 and C7 were examined for groups 1 and 2 data, assuming that 20% of the eggs are contaminated (see above), so that $p = 0.80$, and that 20% of the samples that were positive for SE would not experience growth, so that $v_0 = 0.2$. In actuality, eggs with small number of SE cells or exceptionally young eggs may not experience growth of SE, and for those eggs that did experience growth of SE, the expected value would be larger (based on the condition of there being growth). However, to simplify the calculations, this adjustment was ignored, with the expectation that doing so would not cause a large error. With these assumptions, Equations C6 and C7 were used to solve for $E(k) = 32.22$ and the standard deviation of k, $\text{std}(k) = 65.15$. The parameters for this distribution are given in the following section.

Distribution of SE Cells Initially Contaminating Albumen

To describe variation of SE levels within albumen, a lognormal distribution was applied to the data as described above. Historically, a negative binomial distribution has been used to describe microbial count and density data. For this analysis, however, a lognormal distribution was used as it was necessary to assure positive contamination levels. These issues are discussed below.

To approximate g, a lognormal distribution, it was assumed that ln(k-1) is normally distributed with parameters (ω, ς), where ω is the mean, and ς is the standard deviation. This distribution assures values of k greater than 1. An attempt was made to fit a negative binomial distribution without zeros. While values of parameters were derived which resulted in a skewed distribution, in the simulations to determine the confidence intervals for the values of the parameters, the estimated values were negative approximately 45% of the time. This indicated unstable estimates and thus a lognormal distribution was assumed.

This distribution depends upon estimates of the mean and standard deviation that are usually stable, even with small numbers of samples. From the values of $E(k)$ and std(k) given above, the estimates of the parameters are: ω = 2.60221 and ς = 1.29535. For this distribution, the 95th percentile was 115 cells and the 99th percentile was 276 cells. It should be noted also, that different values of v_0 had a small effect on the estimated values: for v_0 = 0.5, ω = 2.71 and ς = 1.30, and the 99th percentile of k *was* estimated to be 311 cells.

Assigning Uncertainty for Initial SE Contamination in Albumen

Uncertainty of parameter estimates ω and ς were determined by a bootstrap with 5,030 simulations. First, independent random variables n, from a binomial distribution with parameters 9 and 9/266 (Group 1 and 2) and m, from a binomial distribution with parameters 22 and 22/138 (Group 3), were generated. Then n random selections from the 9 results, with replacement, were made. Using the n results, estimates of $E(k)$, var(k), ω and ς were derived, as described above. The results for n less than 3 were deleted, because, if 2 positive results were seen, it is probable that another procedure for estimating the distribution would have been developed.

The distribution of the estimates of ς was negatively skewed due to a small percentage (0.2%) of exceptionally low results. When eliminating these results (which were 3.5 standard deviation units below the mean of all the results), the distribution was nearly symmetrical (skewness = -0.14, from -0.27 with all the results) and kurtosis of 2.65. The distribution of the simulated estimates of ω was negatively skewed with 5 negative results. Eliminating these five results, leaving 4998 remaining bootstrap values, the distribution of ω and ς were estimated: the distribution for ς was nearly symmetric with skewness of -0.16 and a kurtosis of 2.64; and the distribution for ω was negatively skewed with skewness of -0.62 and a kurtosis of 4.5.

To derive a symmetric distribution to use for ω, the quantities, $\varphi(c) = (5- \omega)^c$ were considered for various values of c. When c = 0.6, the skewness of $\varphi(0.6)$ was 0.03 and the kurtosis was 2.19. Thus, for determining the distribution of ω, the distribution of $\varphi(0.6)$ was used. The mean of φ (0.6) is 1.70, which transforms to a value of ω = 2.58, which is close to the estimated value of ω of 2.60 given above.

The standard deviation of the simulated values of $\varphi(0.6)$ is 0.29390; the standard deviation of the simulated values of ς is 0.1425; and the correlation between these two variables is 0.94310. To determine percentiles for these variables that account for the non-zero kurtosis, an Edgeworth

expansion term, $(z^3 - 3z)\kappa_4/24$, where z is a quantile of the standard normal distribution and κ_4 is the kurtosis, was used.

Results and Assumptions for Estimating Initial SE Contamination of Albumen

The results of this section estimate the distribution for the number of SE cells within the albumen (*Ea*) of a contaminated egg at the time of lay. This distribution was dependent on the parameter estimates of the mean and the standard deviation. Further, uncertainty for these estimates was determined. These estimates and the uncertainty will be used in the Exposure Assessment/Risk Characterization section to determine the risk to the consumer of consuming an SE contaminated egg. The assumptions are: 1) it was assumed that at lay, the level, k, of SE cells in *Ea* contaminated eggs is distributed such that $\ln(k\text{-}1)$ is a normal distribution with mean, ω, equal to 2.6022 and standard deviation, ς, equal to 1.2953; and 2) the uncertainties of ω and ς are determined by assuming that $\varphi = (5\text{-}\omega)^{0.6}$ and ς to be nearly normally distributed with means of 1.6997 and 1.2953, and standard deviations of 0.2939 and 0.1425, respectively, with correlation coefficient of 0.9431. Generated standardized values from a bivariate normal distribution with zero means, unit standard deviations, and correlation of 0.9431, say z_1 and z_2, respectively, are

$$z_j' = z_j + \frac{\kappa_4(z_j^3 - 3z_j)}{24}, \quad j = 1, 2 \tag{C8}$$

adjusted by computing

where κ_4 is the kurtosis. For φ, $\kappa_4 = 2.19$ and for ς, $\kappa_4 = 2.64$. These adjusted values, z_j are multiplied by the corresponding standard deviation (0.2939 for φ and 0.1425 for ς) and then added to the corresponding mean values (1.6997 for φ and 1.2953 for ς) to calculate the simulated values of φ and ς. The simulated value of ω is $5\text{-}\varphi^{1/0.6}$.

DISTRIBUTION OF LEVELS OF SE IN YOLK (*EY)* AND VITELLINE MEMBRANE (*EV*) CONTAMINATED EGGS AT TIME OF LAY

The model used to describe the initial levels of SE in the yolk or on the vitelline membrane was guided by a conceptual model describing infection of an egg by an infected hen. SE are capable of contaminating the ovary and oviduct of a hen. Infection of these sites can lead to vertical contamination of the inner yolk contents (*Ey*) or the vitelline membrane (*Ev*). *Ey* and *Ev* events can occur either directly, as indicated above, or by migration of SE from other locations within an SE-infected egg. Depending on the mechanism, some level of SE can be deposited within the yolk or the vitelline membrane. It is the purpose of this section to estimate the level of SE for *Ey* and *Ev* events at the time of lay.

In calculating these estimates, we used data from experimentally inoculated hens. Issues such as separating the contamination of the vitelline membrane vs. internal yolk contents, SE levels

below enumeration and false negatives are considered in interpreting these data. These issues are discussed below.

Estimating the Distribution of the Number of SE Cells in Yolk and on the Vitelline Membrane

Conceptual model

SE can contaminate the yolk or vitelline membrane of an egg by two mechanisms: 1) directly by contamination of the ovary or oviduct and then infection of these compartments; or 2) indirectly by migration from another location within the egg, e.g., the albumen or vitelline membrane.

Direct contamination of the yolk or vitelline membrane could occur early in the formation of the egg by contamination of the ovary and/or contamination of the opening of the oviduct known as the infundibulum. This could be a large or small contamination event. Given the time the egg is within the hen, growth could occur. Therefore, enumeration of SE within egg albumen at the time of lay does not necessarily indicate the initial contamination level.

Indirect contamination of the yolk or vitelline membrane could occur by SE migration from the vitelline membrane or the albumen, respectively. As SE are motile and can move within the egg, after some time SE contaminating the albumen could reside on the vitelline membrane. Further, SE contaminating the vitelline membrane could infect the internal yolk contents. Contamination of these compartments would take time relative to the above route of infection and therefore could demonstrate a lower likelihood of growth.

These mechanisms suggest the SE contamination levels of the yolk or vitelline membrane observed experimentally might demonstrate a multi-modal distribution due to the two routes of infection and their potential for growth. These data are presented below.

Modeling the Distribution of the Number of SE Cells in Yolk and on the Vitelline Membrane

The study used for determining the initial level of SE contamination in yolk was that of Gast and Holt.[5] These authors inoculated hens with SE and collected eggs daily for analysis. Yolk was separated from albumen and the vitelline membrane and the yolk contents were homogenized. To screen for SE positive yolks, 10 mL of yolk/egg from among 874 eggs were enriched in broth. This screen identified 21 yolk SE-positive eggs. From here, a 1 mL aliquot from a 3 mL refrigerated companion sample was analyzed for numbers of SE. Among the 21 SE-positive yolk samples, 18 were reported non-enumerable. That is, the initial SE level within these samples was below the limit of detection. Of the remaining 3 samples that could be enumerated, the levels were 4, 27 and 67 cfu/mL.[6] If the volume of the yolk and vitelline membrane was assumed to be approximately 15 mL, and it was assumed that the distribution of SE cells within the yolk region was uniform, then it could be estimated that there were 1,005 SE cells within the yolk of the 67 cfu/mL egg. This assumption was reasonable because, before the samples were formed, the yolk material was homogenized.

The eggs were collected daily, so that the age of the eggs when the samples were prepared and the contents analyzed are not specifically known. Yolks can become contaminated either directly through ovaries or indirectly through a contamination of the albumen with subsequent migration to the yolk. The latter route suggests the age of the egg could be an important factor in determining the transition from an *Ev* to an *Ey* event. Results presented by Gast and Holt[7] suggest that migration can occur in a relatively short time (10% of *Ev* infected eggs can become *Ey* in 6 hrs), and thus it was presumed that the event could happen while the egg was still being formed in the oviduct or prior to collection. Once in the yolk, rapid growth of SE could commence following a lag period, upon which the contamination would spread throughout the yolk. Consequently, the distribution of the number of SE cells in the yolk among eggs not more than one day old could be highly heterogeneous, possibly, or most likely, multi-modal in nature, due to the different avenues of contamination and growth potential**.**

This expectation seems to be borne out by the results of Gast and Holt.[5] The fact that 86% (18/21) of the known SE positive samples could not be enumerated (suggesting low levels) and 14% (3/21) had high results, suggests a highly skewed, or multi-modal, distribution of SE cell counts within the contaminated yolks of eggs.

For the purposes of analysis and modeling, it can be imagined that the yolk is divided into three regions: 1) the 10-mL volume (= v_1) initial screening sample that was analyzed to determine the presence of SE cells; 2) the 1-mL volume (= v_2) comprising 5 0.2 mL portions taken from a 3 mL sample for enumeration; and 3) the remaining volume (= v_3).

Let v be the total volume, assumed to be 15 mL, and p_j = the probability that a specified cell would be in the subsample of volume v_j, j = 1, 2, 3. Thus, $p_1 = 2/3$, $p_2 = 1/15$, and $p_3 = 4/15$. To compute the probability of SE within each of these regions, let w be the proportion of eggs that are not contaminated, and, for the other eggs, let the cumulative distribution of the number of cells in the egg, x, be $F(x|\theta)$, where θ are parameters. For a given $x > 0$, the probability of the number of cells in the sub-sample, v_j, x_j, j = 1, 2, 3, such that the sum, x, is a multinomial distribution. Furthermore, it was assumed that there was not 100% recovery, but the probability of a specific cell not being recovered is τ. This can be visualized as further dividing the volumes v_1 and v_2 into four subsections: v_{11}, v_{12}, v_{21}, v_{22}, such that $v_{j2} = \tau v_j$ represents the volume for which the cells are not recovered. Thus the actual number of SE cells recovered is x_{11} and x_{21} for the screening and enumeration samples, respectively.

The probability of a positive result on the screening sample is, $Q_+(x) = 1-(1-p_1(1-\tau))^x$. The multinomial structure permits a simple derivation of the probability of detecting y cells in a 1 mL sample, $Q_y(x)$, given that there are x cells in the egg. This event entails the event, A, of y cells in volume v_{21}, which has the binomial probability with parameters x and $p_2(1-\tau)$. However, A includes the disqualifying event B of $x_{11} = 0$ and $x_{21} = y$, thus, the probability of this event must be subtracted from the probability of A. The probability of B is the multinomial event with parameters x, $p_2(1-\tau)$ and $1-(p_1+p2)(1-\tau)$ for the occurrence of $x_{11} = 0$, $x_{21} = y$, and the remaining cells x-y. Thus, performing the subtraction, it is derived that:

$$Q_y(x) = \binom{x}{y}(p_2(1-\tau))^y[(1-p_2(1-\tau))^{x-y} - (1-(p_2+p_1)(1-\tau))^{x-y}] \qquad \text{(C9)}$$

$$Q_y = (1 - w)\int_0^{\infty} Q_y(x)\,dF(x \mid \theta) \qquad \text{(C10)}$$

and the likelihood of the screening test being negative is:

$$Q_{(-)} = w + (1 - w)\int_x^{\infty} (1 - p_1(1 - \tau))^x dF(x \mid \theta) \qquad \text{(C11)}$$

For the purposes of this analysis, it was assumed that the distribution of SE cells in egg yolks can be described as bi-modal: 1) low-valued population: the 18 enumeration-negative results represent one subpopulation of the distribution for which there was no growth. This subpopulation represents *Ev* contaminations; and 2) high-valued population: the 3 enumeration-positive results represent the other subpopulation for which either there was growth or an initial high level. This subpopulation represents *Ey* contaminations.

Below, the distribution and uncertainty for initial contamination of each subpopulation is determined.

Low-Valued Population

Modeling the distribution for the low-valued population

For the low-valued population representing no growth and *Ev* contaminations, assume that *F* is a

$$p(x = k) = \frac{\rho^k / k!}{e^{\rho} - 1} \qquad \text{(C12)}$$

non-zero Poisson distribution with parameter ρ, with probability increment, $dP(k|\rho)$ equal to:

where k = 1, 2, etc. This distribution was selected for convenience, although other distributions such as the logarithmic series distribution could be considered.

Modeling the likelihood of a negative enumeration sample

The likelihood of a negative enumeration sample, i.e., a sample that did not experience growth, was computed by first noting that:

$$\int_{1}^{\infty} Q_0(k)dP(k \mid \rho) = (e^{\rho} - 1)^{-1} \sum_{k=1}^{\infty} \rho^k \frac{(1 - p_{11})^k - (1 - p_{11} - p_{21})^k}{k!} \quad \text{(C13)}$$
$$= (e^{\rho} - 1)^{-1} (e^{\rho(1 - p_{11})} - e^{\rho(1 - p_{11} - p_{21})})$$

where $p_{j1} = p_j(1-\tau)$, so that:

$$Q_o = (1 - w)((e^{\rho} - 1)^{-1}(e^{\rho(1 - p_{11})} - e^{\rho(1 - p_{11} - p_{21})}) \quad \text{(C14)}$$

where 1- w is the percentage of eggs that are contaminated, given that the eggs are from the low-value population or negative.

Modeling the likelihood of a negative screening sample

Similarly, the likelihood that the screening test is negative is:

$$Q_{(-)} = w + (1 - w)(e^{\rho} - 1)^{-1}(e^{\rho(1 - p_{11})} - 1) \quad \text{(C15)}$$

For example, if $\rho = 2$, $\tau = 0.2$, representing an 80% recovery, and $w = 0.95$, then $Q_0 = 3.4\%$.

Modeling the probability of observing 1 or more cells for an enumeration sample

Of further interest was the probability of observing 1 or more cells for an enumeration sample, Q_1. From Equation C12, assuming the non-zero Poisson distribution, or from $1-Q_{(-)}-Q_0$, $Q_{1\ was}$ determined to be:

$$Q_{\geq 1} = \frac{(1 - w)(e^{\rho p_{21}} - 1)(e^{\rho(1 - p_{21})} - e^{\rho(1 - p_{21} - p_{11})})}{e^{\rho} - 1}. \quad \text{(C16)}$$

In the above example, $Q_1 = 0.384\%$, which from 871 samples would imply an expected 3.345 samples for which enumeration would be 1 cell or more. The measure of fit statistic is -2 log-likelihood ratio:

$$\text{‹} = -2[n_{-} ln((n_{-} + n_0)Q_{-} / n_{-}) + n_0 ln((n_{-} + n_0)Q_0 / n_0)] \quad \text{(C17)}$$

where n_{-} is the number of samples that were not positive by the screening test and n_0 is the number of samples with enumeration results of non-detect. The measure of fit statistic has a value of 12, which indicates a lack of fit, when compared to the percentiles of a chi-square with 1 degree of freedom.

The parameter, ρ, was determined by assuming τ is equal to 0.2; setting the parameter Q_1 so that the probability of no enumeration sample with a positive result from the $n_{-} + n_0 = 871$ screening samples is 25%; and minimizing . A solution for ρ and w was obtained by constructing the Lagrangean equations (with one multiplier), setting the derivatives equal to zero, and solving in Mathcad®7. For these constraints, ρ was determined to be 1.3922, $w = 0.9681$ (or about 3% of the eggs were contaminated in the yolk or vitelline membrane) and = 2.773, which is significant with P-value = 0.1, indicating a moderate degree of fit. Uncertainties of these values are determined by considering n_{-} and n_0 as random variables arising from a multinomial distribution with the number parameter equal to 874 samples.

Possibility of Egg Contamination Before or After Lay

Eggs that are members of the low-valued population are assumed to have experienced little growth. This suggests the contamination is not in the yolk, as cells in the yolk would likely leave lag phase quickly and be detected easily. This is important as yolk contaminations can lead to substantial numbers of cells within an egg over a short period of time. Therefore, the need to distinguish between yolk and vitelline membrane contaminations is necessary to establish a realistic estimate of risk. The low-valued eggs might represent small level contaminations in the vitelline membrane due to a migration event from the albumen after the egg was laid. This suggests the contamination is within a location less likely to exhibit rapid growth and could only be within this location from the time of lay to collection (1 day maximum). Eggs with these conditions would experience little growth. However, it is possible the contamination took place before the egg was laid, i.e. direct contamination of the yolk or vitelline membrane and the cells were still in the lag phase. This latter possibility could suggest the contamination was in the

vitelline membrane or the yolk for up to 24 hrs within the hen plus the time before collection. Under these conditions, growth would be more likely and risk would be greater. To evaluate these possibilities and therefore establish risk for the low-valued population, an evaluation of possible growth, or actually, no growth, needs to be examined.

To determine the possibility of no growth for the low-valued population, the probability of no growth for an egg a particular age was first calculated. An initial contamination of 1 to 15 cells was assumed. This was followed by estimating the probability of no growth for a randomly selected egg from a population of eggs. Therefore, assuming the contamination to be uniformly distributed over a time for potential contamination, the possibility for no growth before or after lay was determined.

To determine the probability of no growth for an egg a particular age, equations were developed in Annex E that describe stochastically the growth of SE cells. The temperature of the egg, as a function of time, was assumed to be:

$$T(t) = min(T_i, T_a + (T_i - T_a)e^{-k(t-t_{lay})}) \qquad \text{(C18)}$$

where T_a is the ambient temperature, T_i is the initial egg temperature, t_{lay} is the time that the egg was laid, relative to the time that the contamination entered the yolk or vitelline membrane, and k is the exponential cooling rate (natural logarithm units per day).

If the contamination enters the yolk or vitelline membrane before the egg is laid, then there are values of t such that t-t_{lay} is negative. Assumptions are: 1) T_i = 41.1°C - the body temperature of the hen; 2) T_a = 24.4°C - room temperatures where the hens reside is maintained at about 24.4°C[1]; 3) k = 0.3(24) = 7.2/day based on study of cooling rates of eggs in open stacks;[8] and 4) as temperature changes, the ratio of the lag to the generation time, *Rat*, remains constant.[9] For these calculations, it was assumed that *Rat* = 5.

The model for determining the exponential growth rates, $\mu(t)$, as a function of temperature, is described in annex E. For these calculations it was assumed that the cells are found in the vitelline membrane, and that the egg yolks contain anti-SE antibodies. With these assumptions, the probabilities of no growth by time t for assumed initial number of SE cells, ranging from 1-15, at the beginning of their lag phase at the time of lay (t_{lay} = 0) are given in Table C4. The expected value of the probability of no growth is given in the last row, computed assuming a Poisson distribution with parameter 1.392. The probability of greater than 15 SE cells is very small.

TABLE C4 PROBABILITY OF NO GROWTH BY AGE (DAYS) WHEN THERE ARE AN ASSUMED INITIAL NUMBER OF CELLS THAT CONTAMINATE THE VITELLINE MEMBRANE (*EV* EGGS) AT THE TIME OF LAY. THE EXPECTED VALUE WAS COMPUTED ASSUMING A POISSON DISTRIBUTION WITH PARAMETER 1.392.

	Time (days)					
	0.05	0.10	0.25	0.5	0.8	1.0
Initial Number Cells	p0n	p0n	P0n	p0n	p0n	p0n
1	0.987	0.965	0.910	0.844	0.778	0.737
2	0.974	0.931	0.828	0.713	0.605	0.543
3	0.961	0.898	0.754	0.601	0.471	0.401
4	0.948	0.867	0.686	0.508	0.366	0.295
5	0.935	0.836	0.624	0.429	0.285	0.218
6	0.923	0.807	0.568	0.362	0.221	0.160
7	0.910	0.778	0.517	0.305	0.172	0.118
8	0.898	0.751	0.470	0.258	0.134	0.087
9	0.886	0.724	0.428	0.218	0.104	0.064
10	0.874	0.699	0.389	0.184	0.081	0.047
11	0.863	0.674	0.354	0.155	0.063	0.035
12	0.851	0.651	0.322	0.131	0.049	0.026
13	0.840	0.628	0.293	0.110	0.038	0.019
14	0.829	0.606	0.267	0.093	0.030	0.014
15	0.818	0.584	0.243	0.079	0.023	0.010
Expected Value For No Growth	0.976	0.936	0.843	0.74	0.646	0.592

The results from Table C4 indicate, for instance, that even if there were 10 SE cells in the beginning of lag phase at the time of lay, in a full day, there was a 4.7% chance that no growth would take place. Within half a day, there was an 18.4% chance that no growth would take place.

To determine the probability of no growth for a randomly drawn egg from a population of eggs, let $Eprob0(t|t_b)$ be the expected value of no growth for an egg of age t, given that the contamination occurred at time t_b. From the results in the last row of Table C4 and others for different times not shown, an approximation of $Eprob0(t|t_b)$ of the form $d/(1+bt^c)$, where b, c and d are parameters with values depending on t_b and were estimated from nonlinear regressions. (For the case of Table C4 ($t_b = 0$), $d = 1.01$, $c = 0.92$ and $b = 0.70$). If it is assumed that the ages of these low valued population eggs when sampled were uniformly distributed over 0 to 1 day, then the expected value that there would be no SE growth, $Eprob(t_b)$, in a randomly drawn egg from this population would be the integral of the $d/(1+bt^c)$ from t_b to 1. The results for selected times SE of contamination before the egg is laid are given in Table C5.

TABLE C5 PROBABILITY THERE WOULD BE NO GROWTH OF SE IN A RANDOMLY SELECTED EGG, GIVEN THE CONTAMINATION ENTERED THE VITELLINE MEMBRANE AT *T* DAYS BEFORE THE EGG WAS LAID.

Fraction of Day After Lay	-1	-0.9	-0.75	-0.5	-0.25	0	0.15	0.25	0.5
Probability of No Growth	29.3%	31.9%	36.4%	45.9%	58.6%	75.9%	83.6%	86.6%	92.6%

Figure C1 is a graph of the natural logarithm of the computed probability of no growth versus the time of contamination and a smoothed fourth degree polynomial fit.

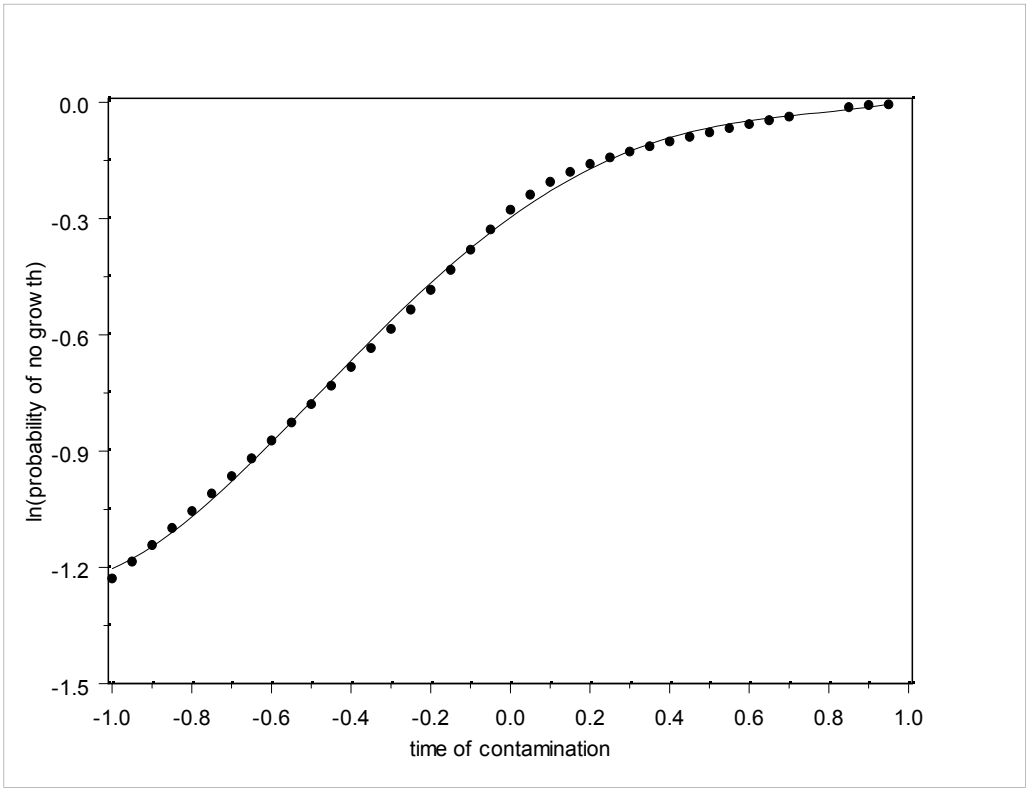

FIGURE C1 COMPUTED NATURAL LOGARITHM OF THE PROBABILITY OF NO GROWTH VERSUS IVEN TIMES OF CONTAMINATION RELATIVE TO WHEN THE EGG IS LAID. THE SMOOTHED LINE IS FOURTH DEGREE POLYNOMIAL LEAST SQUARES REGRESSION.

The above calculation of the probability of no SE growth assumes that the egg is sampled at an age greater than the time of contamination; that is, it assumes that the eggs are sampled at a random time between max(0, t_b) and 1. Eggs sampled before t_b of course would show no growth. Including this latter possibility, the probability of a randomly selected egg of showing no growth is:

$$PNg = \frac{1}{2}\int_{-1}^{1} (max(0, t_b) + (1 - max(0, t_b) E_{prob}(t_b)) dt_b \qquad \text{(C19)}$$

Using the derived polynomial regression, *PNg* was derived to be 71%. This value can be visualized as an average of the conditional probabilities of no growth given that the egg became contaminated before and after lay. The former probability is 48% and the latter is 94%. If the

beginning of the contamination were assumed to be uniformly distributed over the interval (-1, 1)[b], then 75% of the contaminations would be sampled. Of the positive samples, 67% of them would be from contaminations that occurred before lay. Consequently, even with a draw of 8 such eggs, the probability of at least one egg experiencing growth would be high.

This analysis shows that the assumption that a significant percentage of these eggs are contaminated in the yolk before lay is not plausible. Consequently, for these eggs, it was assumed that the contamination occurred after lay and the contamination resides within the vitelline membrane.

RESULTS AND ASSUMPTIONS USED FOR LOW-VALUED POPULATIONS

1) For the low value population, it was assumed that the contamination was located within the vitelline membrane and occurred post-lay.

2) For the low value population, it was assumed that the number of SE cells at the beginning of a contamination was distributed as a Poisson, without zeros (Equation C12), with parameter, ρ, equal to 1.3922.

3) The age of the eggs when the contamination begins in vitelline membrane was distributed uniformly in the interval (0, 1). The assumption that the ages of the eggs when sampled were uniform implies that an expected 50% of the contaminated eggs of this type would not have been sampled.

4) The fraction of low valued positive eggs was estimated as $2(1-w)$, where $w = 0.96813$.

5) Uncertainty of ρ and w was determined by assuming the vector: n_-, n_0, and n_+, is distributed as a multinomial distribution with probability parameters: z_-, z_0 and $z_+ = 1- z_- - z_0$, and number parameter 874. The values from the trinomial distribution were generated by a sequence of two binomial distributions, where the first one generated a value of n_+, labeled m_+ for this discussion, from the binomial (874, 3/874), and then the second one generated a value of n_- from the binomial(874- m_+, 871/874). The strategy is based on the probability law: $P(A, B) = P(A|B)P(B)$, and noting that the conditional probability, $P(A|B)$, and the unconditional probability, $P(B)$, are distributed as binomial distributions. A bootstrap of 1000 simulations was performed, and various transformations of the results were examined to find ones that were nearly normally distributed. For ρ, the transformation $g(\rho) = \ln(1+\ln(\rho))$ was nearly normally distributed with the mean equal to 0.27671, for which the inverse transformation is 1.3755, and the standard deviation equal to 0.17325 (the skewness = 0.062 and the kurtosis = 0.1873). To adjust for the non-zero kurtosis, Equation C8 was used. For w, the transformation, $h(w) = (-\ln(w))^{0.75}$ was nearly normally distributed with the mean equal to 0.077364, for which the inverse transformation is 0.96754, and the standard deviation equal to 0.015473 (the skewness = 0.022 and the kurtosis = 0.007). The correlation of $g(\rho)$ and $h(w)$ is 0.99961.

[b] This time interval indicates 24 hrs prior to egg lay through 24 hrs post-lay. It represents a contamination event within the hen that could have taken place before egg lay through the maximum time the egg could have remained before collection and analysis.

High-Valued Population

For the high-valued population, the age of the sampled egg is critical as the levels observed by Gast and Holt[5] could quickly lead to substantial levels within an egg. For these three eggs (4, 27 and 67 cfu/mL of yolk), it was assumed that there was SE growth taking place when sampled, so that the time that the growth of SE begins is important. Therefore, to identify the initial contamination levels of this subpopulation, the amount of growth that might have occurred for these three eggs needed to be calculated. For this, it was assumed 1) that the SE cells are in the vitelline membrane or yolk; and 2) that at least one of the cells is out of its lag phase.

To begin to identify the initial level of the high-valued results (4, 27 and 67 cfu/mL of yolk), its need to be determined if these three levels could be explained using the Poisson assumption derived above for the initial distribution of the low-valued population. Since the predicted exponential growth rate per day for SE cells in the vitelline membrane at 24.4^{o}C was predicted to be approximately 3.8 $\log_{10}$/day, it is possible to explain the three results by assuming a single cell was out of its lag phase sometime before the egg was sampled at less than or equal to 1 day old. However, using the Poisson distribution would not account for the uncertainty that these three results present and therefore will not be used in the analysis of this subpopulation.

In addition, it is unclear if by the time the eggs were sampled, growth can be assumed to have taken place. It can not be dismissed that the ages of some of these eggs when sampled could be less than just a few hours old and that these eggs represent serious contaminations with high numbers of SE cells, or that at least a few of them were in the exponential growth phase. As discussed in the introduction to this chapter, high levels might occur in contaminations due to the swarming or quorum sensing phenomena (annex A). Alternatively, it is possible initial contaminations of the yolk or vitelline membrane occur early in an egg's development, close to 24 hrs before the egg was laid. This could be due to colonization of the ovary and/or the upper oviduct by SE. By the time the egg was laid there could be a large and growing contamination, even if the initial number of SE cells in the original contamination was not large. Therefore the assumption is possible. The complexities are infinite, and the information is infinitesimal.

The three high results suggest the SE within these yolks could have been growing at the time of collection. If this is true, the time and temperature conditions of the egg would likely impact the expected value of SE. That is, if the eggs were stored differently from the conditions of Gast and Holt,[5] then it would be expected that the measured SE levels would differ from those reported. Consequently, it was necessary to make some adjustments accounting for time and temperature storage conditions.

To make this possible, the high-valued population was modeled in such a way to account for the time and temperature differences potentially experience by an individual egg. The strategy was to back-calculate, to a time, $t(1)$, when it is imagined one cell would have existed out of lag phase, from which the expected subsequent growth would provide a level equal to that at the age that the eggs were sampled under the same conditions.

An Example to Demonstrate Difficulty

It is difficult to determine the possible initial states that could give rise to a specific set of results. An example is given that demonstrates the difficulty of determining an initial set of conditions

by back-calculating possible growth scenarios, and also shows the type of calculations that were done.

In general, there could be N SE cells initially in the contamination, where only some of them actually begin to grow (divide), while the others do not until later. Suppose the age of an egg when sampled was z, and the obtained result was x. It was assumed that, at the time of sampling, SE cells were growing. Let it be assumed that there was a time, t, less than z, when a single SE cell left its lag phase and the measured value of x represents the number of that cell's progeny. The actual distribution of the increase of the number of cells is a geometric distribution,[10] with parameter p, where p is the expected value of X; specifically, the cumulative distribution of the random variable X, representing the increase number of cells, is:

$$G(X) = 1 - (1 - p)^{X+1} \tag{C20}$$

so that the q^{th} percentile, X_q, is:

$$X_q = \frac{ln(1 - q)}{ln(1 - p)} - 1 \tag{C21}$$

If the result of 67 cfu/mL represented the 99th percentile of the distribution given above, then p = 1/261.3, or that the expected value would be 261.3, or, in log base 10 units, 2.42 $\log_{10}$ cfu/mL. If the result represented the 1st percentile, then the expected value would be 5.07 $\log_{10}$ cfu/mL. This latter possibility could happen for the assumed model if a single cell left its lag phase 45 minutes before the egg was laid and the egg was one day old when sampled. However, another possibility is the SE cell left its lag phase 12 hours before the egg was laid and the age of the egg when sampled was about 1.7 hours. If this were to happen then, employing the growth model developed for these risk assessments, it would be expected that the egg at one day old would have about 9 $\log_{10}$ SE cells.

One can go farther and assume that the distribution of levels is lognormal for the population sampled and that the results represent a simple random sample from this distribution. The p^{th} percentile of predicted individual values was determined from:

$$\bar{x} + k_p\left(s\left(1 + \frac{1}{n}\right)\right)^{\frac{1}{2}} \tag{C22}$$

where k_p is p^{th} percentile of the t-distribution with 2 degrees of freedom. Assuming that results represent cfu/mL in 15 mL of the yolk, and the recovery is 80%, then the upper 99.5th percentile of the individual values is approximately 9.5 $\log_{10}$. However, as the ages of the eggs could be between 0 and 1 day old, these results could represent eggs with levels of SE above 9 $\log_{10}$ at 1 day. There are neither data nor scientific theory known that would eliminate this possibility.

ESTIMATING THE INITIAL CONTAMINATION OF THE HIGH-VALUED POPULATION

To estimate the initial contamination of this subpopulation, it was assumed that the eggs, when sampled, were less than one day old. The ages of the eggs when sampled are unknown, so it was assumed that the ages are uniformly distributed over the interval (0, 1). Further, calculations are performed when it was assumed that SE growth takes place in the yolk (*Ey* contaminations), and that the eggs contain antibodies to SE. To distinguish if the three high-valued results are *Ev* or *Ey* contamination, data from the study of Gast and Holt[7] were used. In this study, 29 eggs were found to be contaminated in the yolk or vitelline membrane. Analysis of the internal yolk contents found that only 3 of the 29 were actually yolk contaminations, excluding the membrane. It was assumed that three eggs corresponding to the high values from the study by Gast and Holt[5] are *Ey* contaminations.

The exponential growth rates, $\mu(t)$, as a function of time through the temperature, are described in annex E, and the temperature profile is given in Equation C18. Thus, the expected relative growth between two times, *a* and *b*, is:

$$r(b,a) = \int_a^b \mu(s)ds \,. \tag{C23}$$

The strategy for estimating the initial contamination of the high-valued population is as follows: The value $t(1)$ represents the time that a single SE cell left its lag phase and entered into the exponential phase of growth. If $t(1) < 0$ then, in the risk assessments, the calculation of the number of progeny cells when the egg is laid is made assuming a temperature of 41.1°C. The likelihood of observing *x* (ignoring measurement error), given the age of the egg, *z*, and the time $t = t(1)$ is:

$$Lik(x \mid z, t(1)) = p(z, t(1))(1 - p(z, t(1)))^{x-1} \tag{C24}$$

where $p(z, t) = r(z, t)^{-1}$.

Let $h(t|\alpha, \beta)$ be the beta distribution of $t(1)$ within the interval [-1,1], representing the possible times that a SE cell could contaminate the yolk (lower bound, -1, is based on the assumption that it takes approximately 24 hours for an egg to be formed and laid), where α and β are the parameters, whose values are to be estimated. Since it is assumed that the ages of the eggs when sampled are uniform and observed values of *x* only occur when $t(1) < z$, the unconditional likelihood value, $\text{Lik}(x|\alpha, \beta)$ is:

$$Lik(x \mid \alpha, \beta) = \int_{-1}^{1} \mathrm{h}(\tau \mid \alpha, \beta) \int_{max(0,\tau)}^{1} p(z,\tau)(1 - p(z,\tau)^{x-1} dz d\tau \quad \text{(C25)}$$

Working with this equation to determine α and β is difficult. Thus, a simplified approach for determining values for α and β is used as follows: assume that the distribution of ages of the eggs when sampled is the uniform distribution; then, for a given $t(1)$, the expected value of $\ln(x)$, $TE_x(t(1))$, would be:

$$TE_x(t(1)) = \int_{max(0,t(1))}^{1} \int_{t(1)}^{z} \mu(\mathrm{s}) ds dz \quad \text{(C26)}$$

Using Equation C26, a value of $t(1)$ was estimated, assuming that $TE_x(t(1)) = \ln(x)$. For x = 1256, corresponding to the sample with measured value of 67 cfu/mL, the value of $t(1)$ is 0.101; for the sample with value 27 cfu/mL, the value of $t(1)$ is 0.140; and for the value of 4 cfu/mL, $t(1)) = 0.253$. These values can be considered the observed values of a sample from a population of times, with density $h(t|\alpha, \beta)$. However, the probabilities of samples being observed as growth samples are not the same because, for times $t>0$, the probability of the egg being sampled at an age less than t is 1- t. Thus, the likelihood, $Lik1(t)$, of an observation is:

$$Lik1(t) = h(t \mid \alpha, \beta) min(1, 1 - t) \quad \text{(C27)}$$

reflecting the lower likelihood of being observed, if t is positive. Stirling's formula: $(x/e)^x(2\pi x)^{0.5}$, for approximating the factorial of x was used for large values of α and β. Maximum likelihood estimators (MLE) of α and β were derived using the logarithmic transformations, $\ln(\alpha)$, and $\ln(\beta)$, and the error covariance matrix derived applies to these transformed values.

Calculations were performed on Mathcad® 7. The MLE estimates of $\ln(\alpha)$ and $\ln(\beta)$ are: 4.9065 and 4.6066, with standard errors of 0.8177 and 0.8172, respectively, and a correlation of 0.9957, computed from the Hessian matrix of second derivatives. Figure C2 depicts the curves of the log-likelihood function, fixing $\ln(\alpha)$ and varying $\ln(\beta)$, for values of $\ln(\alpha)$, including the MLE, 4.5, 4.8, 5.0, and 5.3. The curve with the MLE is bolded. As is evident, the loglikelihood function is flat near the MLE estimates, contributing to the relatively large standard errors.

The above approach does not account for the uncertainty that would exit for the values of the parameters due to the uncertainty of the values of the exponential growth rates. The affect of different assumed exponential growth rates would be to translate the times by certain amounts so that the expected value of the distribution would be most affected.

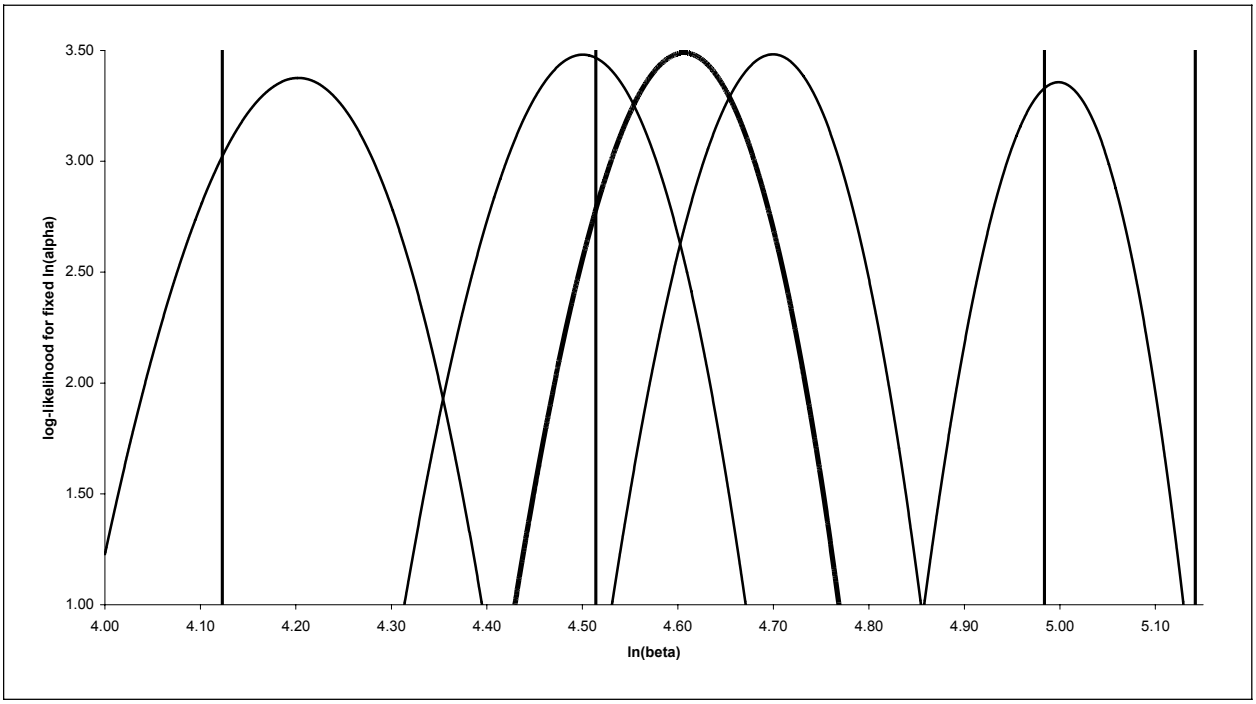

FIGURE C2 PLOTS OF LOGLIKELIHOOD CURVES FOR FIXED VALUES OF LN(∀), WHERE THE X-AXIS IS A RANGE OF LN(*B*) VALUES. THE MIDDLE CURVE (BOLDED) WAS GENERATED WITH THE MLE OF LN(∀); THE MAXIMUM OCCURS FOR THE MLE ESTIMATE OF LN(*A*) = 4.907 AND LN(*B*) = 4.607.

RESULTS AND ASSUMPTIONS USED HIGH-VALUED POPULATIONS

1) It was assumed that the time, $t(1)$, that one SE cell in the interior of the yolk (*Ey* contamination) enters the exponential growth phase was distributed as a beta distribution over the interval [-1, 1] with parameters $\alpha = e^{4.9065}$ and $\beta = e^{4.6066}$.

2) The uncertainty was determined by assuming that $\ln(\alpha)$ and $\ln(\beta)$ were distributed as a t-distribution with means equal to 4.9065 and 4.6066, respectively, with standard errors of 0.8177 and 0.8172, respectively, and a correlation of 0.9957. Because there are three observations, there are 2 degrees of freedom for determining the standard deviation from the mean. Thus, 2 degrees of freedom are associated with these estimates. Formally, this would imply that values of $\ln(\alpha)$ or $\ln(\beta)$ could be negative, or that values of α or β are less than 1, a result that seems counter-intuitive. Consequently, generated values of α or β that are less than 1 are set equal to 1.

3) The age of the eggs when sampled was assumed to be uniform over the interval [0, 1], so that any egg that was contaminated before lay and therefore what have had time to growth to easily detectable levels would have been sampled while if contaminated after lay, would have been sampled with a probability of ½. Thus, there is a 75% chance that the contaminated eggs would have been sampled. Because three contaminations were detected, it is possible that 4 could have existed in the population. Consequently, for these risk assessments, the percentage of eggs that are members of the high value population is (1.33)3/874. To determine the uncertainty of this estimate, the number of positive eggs was assumed to be distributed as 1.33 times a binomial distribution with parameters 3/874 and 874.

REFERENCES

1. Gast RK. Personal communication. 1999.

2. Cogan TA. Personal communication. 2002.

3. Gast RK, Beard CW. Detection and enumeration of *Salmonella* Enteritidis in fresh and stored eggs laid by experimentally infected hens. J Food Prot 1992;55:152-6.

4. Cogan TA, Domingue G, Lappin-Scott HM, Benson CE, Woodward MJ, Humphrey TJ. Growth of *Salmonella enteritidis* in artificially contaminated eggs: the effects of inoculum size and suspending media. Int J Food Microbiol 2001 Oct;70(1-2):131-41.

5. Gast RK, Holt PS. Deposition of phage type 4 and 13a *Salmonella enteritidis* strains in the yolk and albumen of eggs laid by experimentally infected hens. Avian Dis 2000 Jul-2000 Sep;44(3):706-10.

6. Gast RK. Personal communication. 2002.

7. Gast RK, Holt PS. Assessing the frequency and consequences of *Salmonella enteritidis* deposition on the egg yolk membrane. Poult Sci 2001 Jul;80(7):997-1002.

8. Bell DD, Curley RG. Egg cooling rates affected by containers. California Agriculture 1966;2-3.

9. Ross T. Predictive food microbiology models in the meat industry. Meat and Livestock Australia, North Sydney, Australia. 1999.

10. Bharucha-Reid AT. Elements of the theory of the Markov Processes and their applications. New York: McGraw-Hill, 1960.