

Estimation of the POD Function and the LOD of a Qualitative Microbiological Measurement Method

CORDULA WILRICH

Bundesanstalt für Materialforschung und -prüfung, Unter den Eichen 87, D-12200 Berlin, Germany

PETER-THEODOR WILRICH

Institut für Statistik und Ökonometrie, Freie Universität Berlin, Garystrasse 21, D-14195 Berlin, Germany

Qualitative microbiological measurement methods in which the measurement results are either 0 (microorganism not detected) or 1 (microorganism detected) are discussed. The performance of such a measurement method is described by its probability of detection as a function of the contamination (CFU/g or CFU/mL) of the test material, or by the LOD_p , i.e., the contamination that is detected (measurement result 1) with a specified probability p . A complementary log–log model was used to statistically estimate these performance characteristics. An intralaboratory experiment for the detection of *Listeria monocytogenes* in various food matrixes illustrates the method. The estimate of $LOD_{50\%}$ is compared with the Spearman-Kärber method.

Qualitative microbiological measurement methods are those in which the measurement results are either 0 (negative result: microorganism of the defined type not detected) or 1 (positive result: microorganism of the defined type detected).

The probability of detection (*POD*) function $p(d)$ is the functional relationship between the probability p of obtaining a measurement result 1, and the contamination of the test material, quantitatively expressed as the number of CFUs of the microorganism of the defined type per unit of weight or volume, d (CFU/g or CFU/mL).

The LOD_p is the contamination d_p , expressed in CFU/g or CFU/mL, that gives a measurement result 1 with probability p ; p is a specified value, e.g., $p = 50\%$ or $p = 95\%$. The $p(d)$ and d_p values can be estimated on the basis of an intralaboratory experiment or an interlaboratory experiment; however, we deal only with intralaboratory experiments.

In an intralaboratory experiment, the laboratory uses samples of k different matrixes. For matrix i , it performs measurements at q_i different levels j of contamination, d_{ij} ; $j = 1, \dots, q_i$. The number of measurements at level j is n_{ij} . If possible, the numbers of measurements, n_{ij} , should be equal

for all matrixes and all levels: $n_{ij} = n$ for $i = 1, \dots, k$; $j = 1, \dots, q_i$. For each matrix, at least three levels should be included into the experiment, i.e., low, medium, and high contamination.

It is assumed that for each contamination level j the contamination d_{ij} (CFU/g or CFU/mL) of the test material is known with negligible error, either because the test material has been spiked or its contamination has been determined by a colony count.

In order to carry out the n qualitative measurements at a particular level of contamination for a particular matrix, n (almost) identical samples are prepared and measured, and the number of positive results, y_{ij} , is recorded (Table 1).

The Model

The samples necessary for the measurements at a particular contamination level j are made up from a portion of test material for which the contamination d (CFU/g or CFU/mL) is assumed to be known. However, the number of CFUs in each sample will not be equal. In the simplest case, this number is Poisson-distributed with expectation $\mu = A_0 d$, where A_0 is the sample size (g or mL) of the measurement samples. If the measurement method were ideal, the measurement result would be 1 if the sample contained at least 1 CFU, and 0 otherwise. Under the Poisson model, the probability of this event, i.e., the *POD* function of the ideal measurement method, is

$$\begin{aligned} p(d) &= P(\text{at least 1 CFU in the sample} | \mu = A_0 d) \\ &= 1 - P(0 \text{ CFU in the sample} | \mu = A_0 d) \\ &= 1 - \exp(-A_0 d) \end{aligned} \quad (1)$$

Because the sample size A_0 is known, this *POD* function $p(d)$ does not have any unknown parameters. However, every measurement method is not ideal and hence, its *POD* function will deviate from the ideal. Reasons for that are

(1) The distribution of the CFUs in the samples is different from the Poisson distribution because of CFU agglomeration. Agglomeration causes more samples with extreme (very low or very high) numbers of CFUs and, hence, $p(d)$ is smaller than that of the ideal measurement method for each contamination $d > 0$.

(2) Some CFUs in the sample are not detected and, hence, the measurement result erroneously is 0, i.e., the sensitivity of the measurement method is $<100\%$. Hence, $p(d)$ is smaller

Table 1. Results of an intralaboratory experiment for the detection of *L. monocytogenes* in $k = 5$ different food matrixes^a

Matrix number i	Food matrix	Inoculation series j	Level of inoculum d_{ij}	No. of tests	
				Inoculated n_{ij}	Positive y_{ij}
1	Pasteurized milk	1	0.0112	6	1
1	Pasteurized milk	2	0.0224	6	2
1	Pasteurized milk	3	0.0448	6	4
1	Pasteurized milk	4	0.0672	6	4
1	Pasteurized milk	5	0.1416	6	6
2	Rillettes	1	0.0156	6	1
2	Rillettes	2	0.0308	6	4
2	Rillettes	3	0.0620	6	4
2	Rillettes	4	0.0928	6	5
2	Rillettes	5	0.0980	6	6
3	Fish	1	0.0144	6	1
3	Fish	2	0.0292	6	5
3	Fish	3	0.0580	6	4
3	Fish	4	0.0872	6	6
4	Frozen cooked vegetables	1	0.0108	6	2
4	Frozen cooked vegetables	2	0.0216	6	4
4	Frozen cooked vegetables	3	0.0432	6	4
4	Frozen cooked vegetables	4	0.0648	6	6
5	Process water	1	0.0200	6	3
5	Process water	2	0.0400	6	2
5	Process water	3	0.0800	6	5
5	Process water	4	0.1200	6	6

^a The sample size was $A_0 = 25$ g.

than that of the ideal measurement method for each contamination $d > 0$.

(3) When a sample with no contamination is measured, the measurement erroneously is 1, because the measurement method reacts positively to microorganisms other than that of the defined type, i.e., the specificity of the measurement method is <100%. Hence, $p(d)$ is larger than that of the ideal measurement method for each contamination d . However, in a planned experiment, we can assume that no microorganisms other than those of the defined type exist in the test material, or that its microflora is noncompetitive so that $p(0) = 0$.

It can be conjectured that these effects are matrix-dependent. Therefore, we introduce a matrix effect F_i of the matrix i into our model:

$$p(d) = 1 - \exp(-A_0 F_i d) \quad (2)$$

The $p(d)$ value is monotonically increasing from $p(0) = 0$ to $p(d) \rightarrow 1$ for $d \rightarrow \infty$. If $F_i < 1$ ($F_i > 1$), then $p(d)$ of matrix i is smaller (larger) than that of the ideal measurement method for each contamination $d > 0$, i.e., the *POD* curve lies below (above) that of the ideal measurement method (Figure 1).

Generally, in the model of Equation 2, *POD* curves for different F_i do not intersect.

The result X of a measurement at the contamination level d is 1 with probability p and 0 with probability $1 - p$, where p is given by Equation 2, i.e., X is Bernoulli-distributed with expectation $\mu = p$. The number of positive measurement results, i.e., 1, in a series of n independent measurements at this level, $Y = \sum_{l=1}^n X_l$, is binomially distributed with parameters p and n . We can write the model Equation 2 as

$$\begin{aligned} p &= 1 - \exp(-\exp(\ln A_0 + \ln F_i + \ln d)) \\ &= 1 - \exp(-\exp(a_0 + f_i + \ln d)) \\ &= 1 - \exp(-\exp(\eta)) \end{aligned} \quad (3)$$

with $a_0 = \ln A_0$, $f_i = \ln F_i$ and

$$\eta = a_0 + f_i + \ln d \quad (4)$$

η is linearly dependent on $\ln d$ and is related to the expectation p of X by

$$\eta = \ln(-\ln(1 - p)) \quad (5)$$

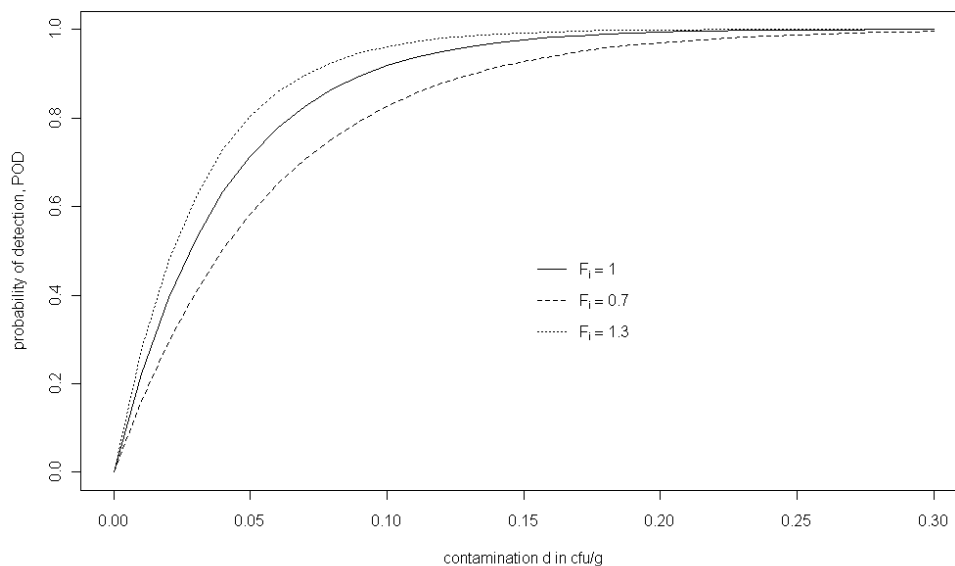


Figure 1. The POD curves for $F_i = 1$ (ideal measurement method), $F_i = 0.7$ and $F_i = 1.3$.

This model is known as the complementary log–log model (cloglog model) because it relates the complementary log–log function of p linearly to $\ln d$. It is a special case of the Generalized Linear Model (1). Equation 5 is a nonlinear transformation of the parameter p into the parameter η that can be linearly predicted from the logarithm of the contamination, $\ln d$. Equation 5 is called the link function and Equation 4 is called the linear predictor. The scope of this experiment and its statistical analysis is to estimate the unknown parameter F_i or f_i in the POD function $p(d)$ for each matrix i .

Estimation of the POD Function

We estimate the k unknown parameters $f_i; i = 1, \dots, k$ using the Maximum Likelihood method. The estimates \hat{f}_i are the roots of the equations

$$\sum_{j=1}^{q_i} \left(\frac{y_j d_{ij}}{\exp(\exp(\hat{\eta}_{ij})) - 1} - (n_{ij} - y_j) d_{ij} \right) = 0; i = 1, \dots, k \quad (6)$$

with $\exp(\exp(\hat{\eta}_{ij})) = \exp(\exp(a_0 + \hat{f}_i + \ln d_{ij})) = \exp(A_0 \hat{F}_i d_{ij})$. The expectation of \hat{f}_i is $E(\hat{f}_i) = f_i$ and its SD is

$$\sigma_{\hat{f}_i} = \frac{1}{\sqrt{\sum_{j=1}^{q_i} \frac{n_{ij} (A_0 F_i d_{ij})^2}{\exp(A_0 F_i d_{ij}) - 1}}} \quad (7)$$

$\sigma_{\hat{f}_i}$ is estimated by

$$s_{\hat{f}_i} = \frac{1}{\sqrt{\sum_{j=1}^{q_i} \frac{n_{ij} (A_0 \hat{F}_i d_{ij})^2}{\exp(A_0 \hat{F}_i d_{ij}) - 1}}} \quad (8)$$

A derivation of Equations 6 and 7 is given in the Annex. The Maximum Likelihood estimator \hat{f}_i is approximately normally distributed and, hence, an approximate confidence interval for f_i with confidence level $1 - \alpha$ is $\hat{f}_i \pm z_{1-\alpha/2} \cdot s_{\hat{f}_i}$, where $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standardized normal distribution. For $\alpha = 0.05$, $z_{1-\alpha/2} = z_{0.975} = 1.96 \approx 2$.

For each matrix i , the estimated POD function is fully determined by the estimate $\hat{F}_i = \exp(\hat{f}_i)$,

$$\hat{p}_i(d) = 1 - \exp(-A_0 \hat{F}_i d) \quad (9)$$

and by inserting the confidence limits for f_i into Equation 9 we get a confidence band for the POD function with the limiting curves

$$p_i(d)_L = 1 - \exp(-A_0 d \exp(\hat{f}_i + 2s_{\hat{f}_i})) = 1 - \exp(-A_0 \hat{F}_i d / \hat{K}_i) = \hat{p}_i(d / \hat{K}_i), \quad (10)$$

$$p_i(d)_U = 1 - \exp(-A_0 d \exp(\hat{f}_i - 2s_{\hat{f}_i})) = 1 - \exp(-A_0 \hat{F}_i d \hat{K}_i) = \hat{p}_i(d \cdot \hat{K}_i)$$

with

$$\hat{K}_i = \exp(2s_{\hat{f}_i}), \quad (11)$$

(Figure 2).

Estimation of the LOD_p

The LOD_p is the contamination d_p , expressed in CFU/g or CFU/mL, that gives a measurement result 1 with probability p ; p is a specified value, e.g., $p = 50\%$ or $p = 95\%$. The value $\hat{d}_{p,i}$ for which the estimated POD is equal to p is an estimate of the LOD_p:

$$\hat{d}_{p,i} = -\frac{\ln(1-p)}{A_0 \hat{F}_i} \quad (12)$$

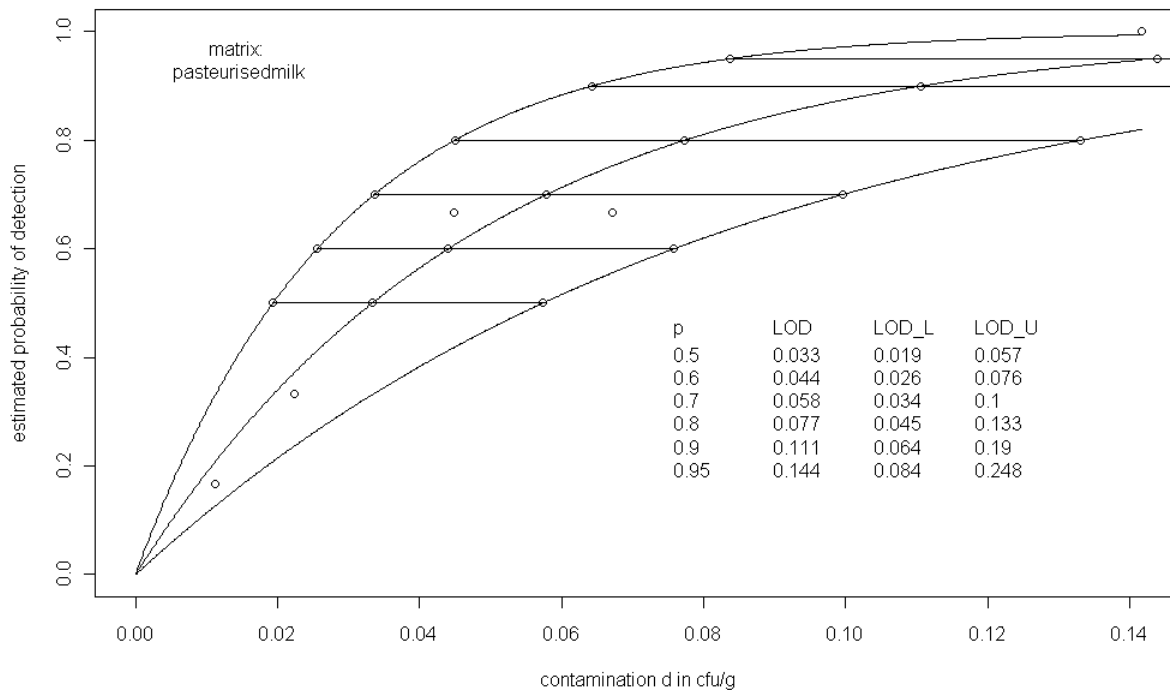


Figure 2. The estimated *POD* curve for pasteurized milk and the 95% confidence band for the *POD*. The experimental results are shown as points (d_{ij}, \hat{p}_{ij}) where $\hat{p}_{ij} = y_{ij} / n_{ij}$ are the observed relative frequencies of detections. The estimated LODs d_p and their 95% confidence intervals for various values of p have been added.

Analogously, the points where the lower/upper limiting curve of the confidence band according to Equation 10 is equal to p is an upper/lower confidence limit for the LOD,

$$\begin{aligned} d_{p,i,L} &= -\hat{d}_{p,i} / \hat{K}_i \\ d_{p,i,U} &= -\hat{d}_{p,i} \cdot \hat{K}_i \end{aligned} \quad (13)$$

where \hat{K}_i is defined in Equation 11.

Figure 2 shows horizontal lines for some values of p . Their intersections with the estimated *POD* curve and its confidence band give the LODs and their confidence limits.

One observes that the values $\hat{d}_{p,i}$ for different p differ only in the value $\ln(1 - p)$. For example,

$$\begin{aligned} \hat{d}_{0.5,i} &= \frac{\ln 0.5}{A_0 \hat{F}_i} \\ \hat{d}_{0.95,i} &= \frac{\ln 0.05}{A_0 \hat{F}_i} \\ \Rightarrow \hat{d}_{0.95,i} &= \frac{\ln 0.05}{\ln 0.5} \hat{d}_{0.5,i} = 4.32 \cdot \hat{d}_{0.5,i} \end{aligned} \quad (14)$$

Further, one observes that the relative uncertainty of the estimated \hat{d}_p , expressed by its confidence limits, is equal for each p , because the confidence limits are derived by multiplication/division with the factor \hat{K}_i , and this does not depend on p . It depends only on the experimental design and the results of the experiment. In order to get a more precise

estimation of d_p , i.e., a smaller SD $s_{\hat{f}_i}$, the total number of measurements (at all contamination levels) has to be increased. If a contamination level is added, it should be equal or close to the value $A_0 d = 1.59$, for which the *POD* under the Poisson model would be $1 - \exp(-A_0 d) = 1 - \exp(-1.59) = 0.80$. At this contamination level, the contribution to the sum in the denominator of $s_{\hat{f}_i}$ (Equation 8) is largest and, hence, the contribution to $s_{\hat{f}_i}$ smallest.

Test of Existence of a Matrix Effect

For a particular matrix i , the value $f_i = 1n F_i$ or its estimate $\hat{f}_i = \ln \hat{F}_i$ describes the deviation of the *POD* curve from the ideal *POD* curve, i.e., the one without measurement error for which $f_i = 0$. Under the null hypothesis H_0 that $f_i = 0$, the expectation of \hat{f}_i is 0, and its standard deviation according to Equation 7 is

$$\sigma_{0, \hat{f}_i} = \frac{1}{\sqrt{\sum_{j=1}^{q_i} \frac{n_{ij} (A_0 d_{ij})^2}{\exp(A_0 d_{ij}) - 1}}} \quad (15)$$

Hence, in order to test the null hypothesis $H_0: f_i = 0$ against the alternative hypothesis $H_1: f_i \neq 0$ at the significance level α , we compare

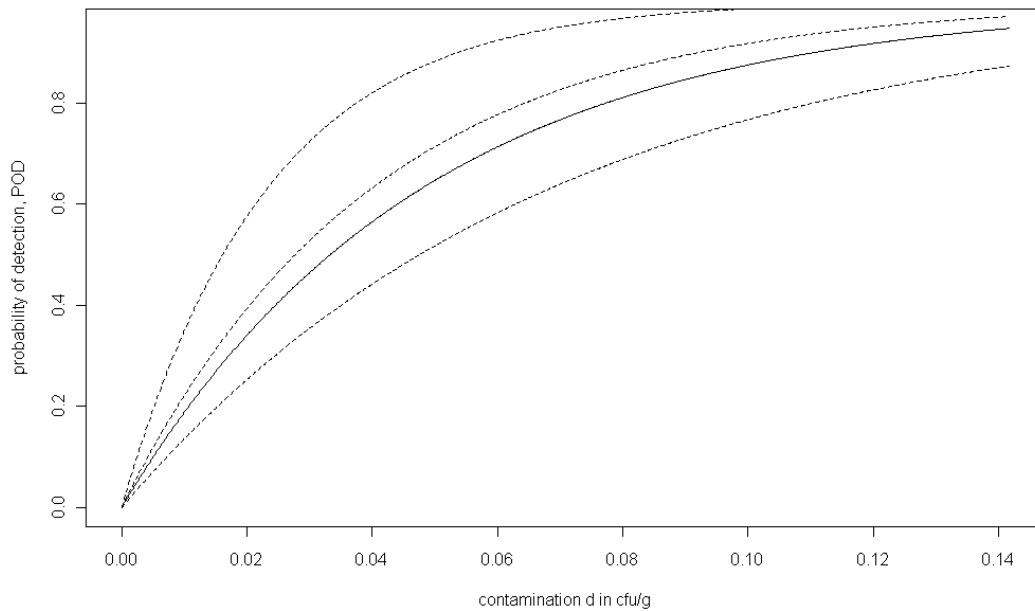


Figure 3. The *POD* curve and the 95% band of variation of the estimated *POD* curve of the ideal measurement method (dashed). The estimated *POD* curve for pasteurized milk (solid) lies inside the 95% band.

$$|z_i| = \left(|\hat{f}_i| \right) \sqrt{\sum_{j=1}^{q_i} \frac{n_{ij} (A_0 d_{ij})^2}{\exp(A_0 d_{ij}) - 1}} \quad (16)$$

with $z_{1-\alpha/2}$, the $(1 - \alpha/2)$ quantile of the standardized normal distribution. For $\alpha = 0.05$, $z_{1-\alpha/2} = z_{0.975} = 1.96 \approx 2$. If $|z_i| \leq z_{1-\alpha/2}$ the null hypothesis $f_i = 0$ is not rejected, otherwise it is rejected.

Under the null hypothesis $H_0 : f_i = 0$, we expect the observed value \hat{f}_i within the interval $0 \pm z_{1-\alpha/2} \cdot \sigma_{0,\hat{f}_i}$ with probability 95% (approximately). Equivalently, we can prepare a graph of the 95% random variation band for the estimated *POD* curve under the null hypothesis $f_i = 0$ with the limiting curves $1 - \exp(-A_0 d \exp(z_{1-\alpha/2} \cdot \sigma_{0,\hat{f}_i}))$ and $1 - \exp(-A_0 d \exp(-z_{1-\alpha/2} \cdot \sigma_{0,\hat{f}_i}))$. If the estimated *POD* curve according to Equation 9 lies inside this band, the null hypothesis $f_i = 0$ is not rejected, otherwise it is rejected (Figure 3).

Finally, it should be mentioned that the value of F_i or its estimate \hat{F}_i has an illustrative interpretation. Suppose we have run an experiment with samples of size A_0 and have found an estimated value of \hat{F}_i . The term $A_0 \hat{F}_i$ in Equation 9 defines a fictitious sample size $A_0^* = A_0 \hat{F}_i$ for which the *POD* curve would be identical to the ideal one. For example, if $A_0 = 25$ g and $\hat{F}_i = 0.8$, then $A_0^* = 25 \cdot 0.8 = 20$ g, i.e., the *POD* curve of the measurement system is equivalent to that for the measurement of samples of size 20 g with the ideal measurement system, i.e., it is worse than the ideal measurement system. On the other hand, if $\hat{F}_i = 1.2$, then the *POD* curve of the measurement system is equivalent to that for the measurement of samples of size 30 g with the ideal measurement system, i.e., it is better than the ideal

measurement system. If, in the latter case of $\hat{F}_i > 1$, the null hypothesis $H_0 : f_i = 0$ resp. $F_i = 1$ was not rejected, this can be explained by random variations in the data; if it was rejected, it needs a careful discussion with the microbiologists, e.g., on lack of specificity.

An Example

Table 1 shows the measurement results of an intralaboratory experiment for the detection of *Listeria monocytogenes* in $k = 5$ different food matrixes. These data were obtained in a study conducted to validate a commercial test kit, anonymously referenced as method J, in the frame of the AFNOR validation scheme (AFNOR Certification, Saint-Denis, France). For each matrix, $n = 6$ qualitative measurements were performed at four or five different contamination levels d , expressed in CFU/g. The sample size was $A_0 = 25$ g. Table 2 presents the results of the statistical analysis for all five matrixes and for the combined data of all matrixes (last line). The first part of Table 2 shows the basic results, \hat{F}_i and $s_{\hat{f}_i}$, the second and third parts give the $\text{LOD}_{50\%}$ and the $\text{LOD}_{95\%}$ and their confidence limits, respectively, and the last column shows the test statistic $|z_i|$ for the test of existence of a matrix effect.

The values \hat{F}_i vary between 0.833 and 1.594. For all matrixes, the null hypothesis $H_0 : f_i = 0$ of no matrix effect is not rejected at the significance level $\alpha = 0.05$.

This test is applied to all five matrixes simultaneously. The probability of erroneously rejecting at least one of the null hypotheses is $1 - (1 - \alpha)^5 = 1 - 0.95^5 = 0.226$, i.e., larger than the significance level 0.05. In order to have it equal to 0.05, the individual significance level has to be decreased to $\alpha = 1 - \sqrt[5]{1 - 0.05} \approx 0.05 / 5 = 0.010$. This is known as the

Table 2. Results of the statistical analysis of the intralaboratory experiment for the detection of *L. monocytogenes*

<i>i</i>	Food matrix	Matrix effect, \hat{F}_i	SD of log matrix effect, $s_{\hat{F}_i}$	LOD _{50%}			LOD _{95%}			Test statistic for matrix effect, $\frac{ z }{ z_c }$
				Detection limit, $\hat{d}_{0.5,i}$	Lower conf. limit, $\hat{d}_{0.5,i,L}$	Upper conf. limit, $\hat{d}_{0.5,i,U}$	Detection limit, $\hat{d}_{0.95,i}$	Lower conf. limit, $\hat{d}_{0.95,i,L}$	Upper conf. limit, $\hat{d}_{0.95,i,U}$	
1	Pasteurized milk	0.833	0.272	0.033	0.019	0.057	0.144	0.084	0.248	0.679
2	Rillettes	0.932	0.251	0.030	0.018	0.049	0.129	0.078	0.213	0.279
3	Fish	1.213	0.283	0.023	0.013	0.040	0.099	0.056	0.174	0.676
4	Frozen cooked vegetables	1.594	0.283	0.017	0.010	0.031	0.075	0.043	0.132	1.571
5	Process water	0.886	0.283	0.031	0.018	0.055	0.135	0.077	0.238	0.426
6	Combined	1.034	0.123	0.027	0.021	0.034	0.116	0.091	0.148	0.267

Bonferroni correction (2). The critical value for this significance level is $z_{1-0.01/2} = 2.57$. Because none of the null hypotheses had to be rejected with the smaller critical value 2, the test result with the adjusted critical value is not different from the former one.

Figure 2 shows the estimated *POD* curve for pasteurized milk and the 95% confidence band for the *POD*. The estimated LODs, \hat{d}_p , and their 95% confidence intervals for various values of *p* have been added.

Figure 3 illustrates the test of existence of a matrix effect for pasteurized milk. The dashed line in the middle is the *POD* curve of the ideal measurement method, $p(d) = 1 - \exp(-A_0d)$, and the two dashed curves above and below it are the limits of the 95% band of variation under the null hypothesis $F_i = 1, p_{0,i}(d)_L = 1 - \exp(-A_0d \exp(2\sigma_{0,\hat{F}_i}))$ and $p_{0,i}(d)_U = 1 - \exp(-A_0d / \exp(2\sigma_{0,\hat{F}_i}))$. The solid curve is the estimated *POD* curve for pasteurized milk ($\hat{F}_i = 0.833$). Because it is inside the band of 95% variation, the null hypothesis is not rejected; for pasteurized milk, the measurement method can be considered as being ideal.

Figure 4 corresponds to Figure 3; however, the estimated *POD* curves for all *k* = 5 matrixes included in the experiment have been added. In our experiment, σ_{0,\hat{F}_i} is not equal for all matrixes because the number *q_i* of levels of contamination and the values *d_{ij}* of the contaminations are different; hence, the 95% bands of variation would be different for the matrixes. Therefore, we have used the largest σ_{0,\hat{F}_i} for the calculation of the 95% band of variation. All five estimated *POD* curves lie inside the 95% band of variation, i.e., the measurement method can be considered as being ideal for all matrixes included in the experiment.

Because the matrix effects of all *k* = 5 matrixes are not significant, we assume that they do not exist and combine the measurement results of all *k* = 5 matrixes, i.e., we carry out the statistical analysis for the whole data set with the matrix information neglected (of course, this statistical analysis can be done with the same easy algorithm as all the others for particular matrixes). We find $\hat{F}_{combined} = 1.034$ and $s_{\hat{F}_{combined}} = 0.123$ (Table 2, last line). Figure 5 shows the *POD* curve and the 95% band of variation of the estimated combined *POD* curve of the ideal measurement method. The estimated *POD* curve for all matrixes combined lies inside the 95% band of variation. The measurement method can be considered as ideal for all matrixes included in the experiment.

A Comparison with the Spearman-Kaerber Method

The Spearman-Kaerber method (3–5) estimates the *LOD*_{50%} by a weighted average of the data. It requires a smallest contamination level for which the fraction of detection results, *y/n*, is 0, and a largest contamination level for which the fraction of detection results, *y/n*, is 1. For its computation, we used the measurement results for pasteurized milk in Table 1. Because the required smallest contamination level with fraction 0 of detection results is missing, we amended the data set by one (*n* = 1) pseudo measurement result *y* = 0 at the very small contamination level *d* = 0.007.

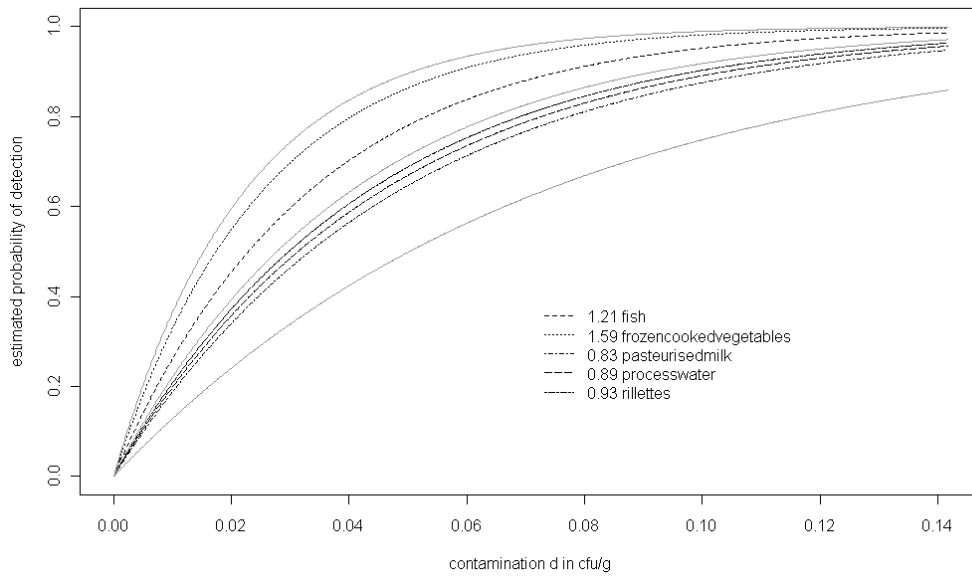


Figure 4. The POD curve and the 95% band of variation of the estimated POD curve of the ideal measurement method (solid). The estimated POD curves for all $k = 5$ matrixes lie inside the 95% band.

Because the fraction of detection results at the largest contamination level is 1, we do not need to amend the data set by a larger contamination level. Table 3 shows the amended data set (the subscript i for the matrix is omitted because we deal with only one matrix).

In the following, the number of contamination levels with eventually one or two amended contamination levels is k ; in our example we now have $k = 6$. We consider the logarithms (to the base 10) of the contaminations, $\lg d_j$, as class limits and the ratios $\hat{p}_j = y_j / n_j$ as cumulated relative frequencies of detections of a classified frequency distribution. The class midpoints of the classes are

$$m_j = \frac{\lg d_{j-1} + \lg d_j}{2}; j = 2, \dots, k \quad (17)$$

and the relative class frequencies are

$$f_j = \hat{p}_j - \hat{p}_{j-1}; j = 2, \dots, k \quad (18)$$

These values are shown in Table 3.

The average of the logarithms of detected contaminations is

$$\hat{\mu} = \sum_{j=2}^k f_j m_j \quad (19)$$

Under the assumption that the distribution of the logarithms of the detected contaminations is symmetric, this is not only an estimator of the theoretical mean but also of the theoretical median, the logarithm of $LOD_{50\%}$. Cornfield and Mantel (6) have given the variance of this estimator,

$$\sigma_{\hat{\mu}}^2 = \sum_{j=2}^k \frac{(p_j - p_{j-1})(1 - p_j + p_{j-1})}{n_j - n_{j-1}} \left[\frac{\lg d_j - \lg d_{j-1}}{2} \right]^2 \quad (20)$$

where $p_j - p_{j-1}$ are the theoretical class frequencies. An estimate of $\sigma_{\hat{\mu}}^2$ is

$$\hat{\sigma}_{\hat{\mu}}^2 = \sum_{j=2}^k \frac{f_j (1 - f_j)}{n_j - n_{j-1}} \left[\frac{\lg d_j - \lg d_{j-1}}{2} \right]^2 \quad (21)$$

$10^{\hat{\mu}}$ is an estimator of $LOD_{50\%}$. Assuming an (approximate) normal distribution of the estimator $\hat{\mu}$, we obtain the interval $\mu_L^U = \hat{\mu} \pm 2\hat{\sigma}_{\hat{\mu}}$ as an approximate 95% confidence interval for the logarithm of $LOD_{50\%}$; $10^{\mu_L^U}$ is an approximate confidence interval for $LOD_{50\%}$.

With the data of Table 3, we obtain $10^{\hat{\mu}} = 0.033$ as estimate of $LOD_{50\%}$ and $10^{\mu_L^U} = [0.020, 0.054]$ as an approximate 95% confidence interval for $LOD_{50\%}$. The estimate is identical to the one obtained in the *Estimation of the LOD_p* section above, $\hat{d}_{50\%} = 0.033$. The confidence interval is slightly smaller than that obtained with that method (0.019, 0.057).

In order to compare the Spearman-Kaerber estimator with the estimator in that section, we run the following simulation experiment. In Table 3, the numbers of detections (without the pseudo measurement result in row 0) are $y_j = 1, 2, 4, 4, 6$. If the measurement method were ideal, the numbers y_j of detections would follow binomial distributions with sample sizes $n_j = 6$ and detection probabilities $p_j = 1 - \exp(-25d_j)$. We performed $N = 10\,000$ simulations with the numbers y_j of detections chosen randomly from these binomial distributions.

Figure 6 shows the *POD* function of the ideal measurement method and the limits of the 95% interval of variation of the *POD* functions estimated from the simulation experiment (solid lines). The dashed lines are the corresponding simulation results, i.e., the average simulated *POD* function and the 2.5% and the 95% quantile of the $N = 10\,000$

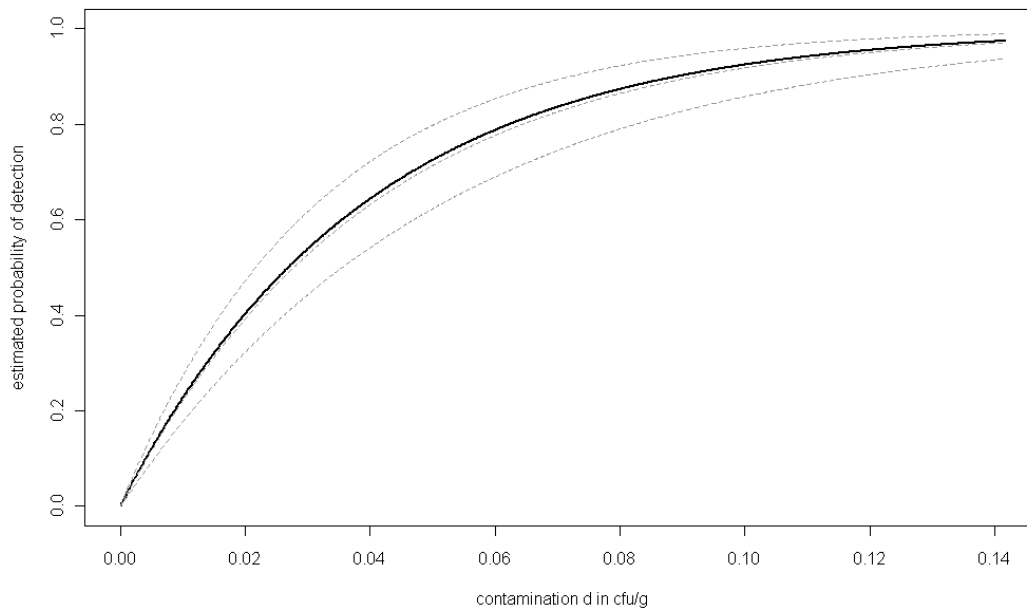


Figure 5. The *POD* curve and the 95% band of variation of the estimated combined *POD* curve of the ideal measurement method (dashed). The estimated *POD* curve for all matrixes combined (solid) lies inside the 95% band.

simulated *POD* functions. The agreement between theoretical and simulated results is almost ideal. The true $LOD_{50\%} = -\log 0.5/25 = 0.0277$. With the estimation of $LOD_{\hat{p}}$ applied to the $N = 10\,000$ simulated measurement results, we find an average estimate $\hat{d}_{50\%} = 0.0284$; the 95% confidence intervals for the $LOD_{50\%}$ have an average length of 0.0326, and 95.2% of them cover the true $LOD_{50\%}$. With the Spearman-Kaerber method we find an average estimate $\hat{LOD}_{50\%} = 0.0257$; the 95% confidence intervals for the $LOD_{50\%}$ have an average length of 0.0281, and 89.1% of them cover the true $LOD_{50\%}$. Other simulations gave similar results. The average length of the Spearman-Kaerber confidence interval for the $LOD_{50\%}$ is smaller than that of the estimation of $LOD_{\hat{p}}$, but its actual confidence level is <90% instead of the expected 95%. These

differences might be due to the fact that pseudo measurements have to be added in order to apply the Spearman-Kaerber method if the number of detections at the lowest concentration level is not 0 or the number of detections at the largest concentration level is not 100%.

There are many reasons to prefer the estimation of $LOD_{\hat{p}}$ method: It not only allows the estimation of $LOD_{50\%}$ but also the values of LOD for other values of p , and it allows the estimation of the *POD* function; it does not need pseudo measurement results in order to be applied; it does not cause difficulties if the series of proportions of detections obtained at increasing contamination levels is not monotonically increasing; and its actual confidence level of the confidence interval for $LOD_{50\%}$ is equal to the chosen confidence level of 95%.

Table 3. Amended data set for *L. monocytogenes* in pasteurized milk^a

Inoculation series j	Level of inoculum, d_j (CFU/g)	No. of tests		Cumulated relative frequency, \hat{p}_j	Log of inoculation level, $\lg d_j$	Class midpoint, m_j	Relative class frequency, f_j
		Inoculated, n_j	Positive, y_j				
1	0.0070	1	0	0.000	-2.155	—	—
2	0.0112	6	1	0.167	-1.951	-2.053	0.167
3	0.0224	6	2	0.333	-1.650	-1.800	0.167
4	0.0448	6	4	0.667	-1.349	-1.499	0.333
5	0.0672	6	4	0.667	-1.173	-1.261	0.000
6	0.1416	6	6	1.000	-0.849	-1.011	0.333

^a The sample size was $A_0 = 25$ g.

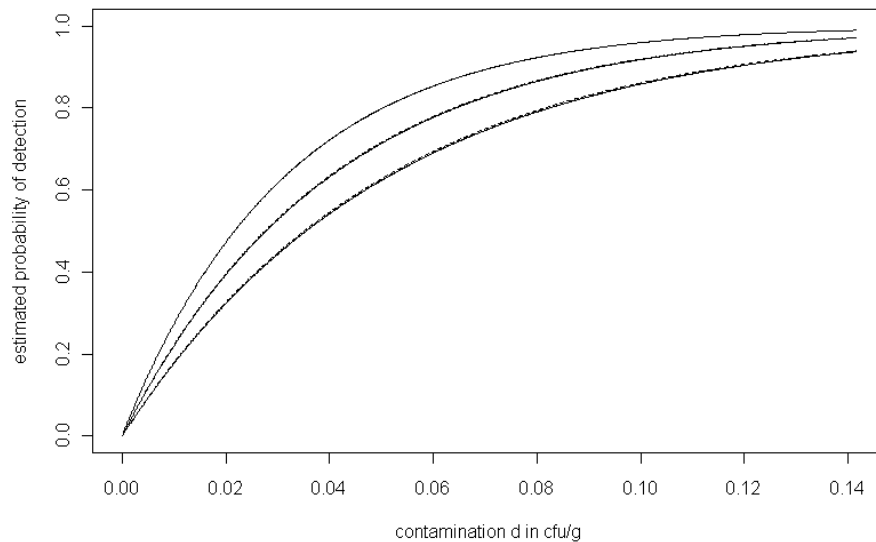


Figure 6. The *POD* curve and the 95% band of variation of the estimated *POD* curve of the ideal measurement method (solid) and the corresponding curves derived from the simulation experiment with $N = 10\,000$ simulation runs (dashed).

Conclusions

The estimation of the *POD* function and the LOD_p of a qualitative microbiological measurement method is based on a complementary log–log model, which is the natural model for such measurements. The estimation includes confidence intervals for the estimated parameters. Additionally, a test of existence of a matrix effect is presented. The estimate of $LOD_{50\%}$ compares more than favorably with the Spearman-Kärber estimate. Based on the results of an intralaboratory experiment for various matrixes, the statistical analysis can be carried out by application of the EXCEL program *PODLOD.xls* that can be downloaded from <http://www.wiwiss.fuberlin.de/institute/iso/mitarbeiter/wilrich/index.html>.

Acknowledgments

We thank three anonymous reviewers for their valuable comments and suggestions.

References

- (1) McCullagh, P., & Nelder, J.A. (1982) *Generalized Linear Models*, Chapman and Hall, London, UK, and New York, NY
- (2) Bonferroni, C.E. (1935) in *Studi in Onore del Professore Salvatore Ortu Carboni*, Rome, Italy, pp 13–60
- (3) Spearman, C. (1908) *Br. J. Psychol.* **2**, 227–242
- (4) Kärber, G. (1931) *Arch. Exp. Pathol. Pharmacol.* **162**, 480–487
- (5) Kotz, S., Read, C.B., Balakrishnan, N., & Vidakovic, B. (Eds) (2006) *Encyclopedia of Statistical Sciences*, Vol. IV, John Wiley & Sons Inc., Hoboken, NJ, pp 354–357

- (6) Cornfield, J., & Mantel, N. (1950) *J. Am. Stat. Assoc.* **45**, 181–210
- (7) Mood, A.M., Graybill, F.A., & Boes, D.C. (1974) *Introduction to the Theory of Statistics*, McGraw-Hill Kogakusha Ltd, Tokyo, Japan

Annex

We estimated the k unknown parameters f_i ; $i = 1, \dots, k$ of the model of Equations 3 and 4 with the Maximum Likelihood method. To do that, we looked at the probability of observing y_{ij} results 1 among n_{ij} independent measurements at contamination level d_{ij} at matrix i ,

$$f(y_{ij}; n_{ij}, d_{ij}, a_0, f_i) = \binom{n_{ij}}{y_{ij}} (1 - p_{ij})^{y_{ij}} p_{ij}^{n_{ij} - y_{ij}} \quad (22)$$

with

$$p_{ij} = 1 - \exp(-\exp(\eta_{ij})) = 1 - \exp(-\exp(a_0 + f_i + \ln d_{ij})) \quad (23)$$

Because the measurements at different levels and different matrixes are independent, the probability of observing the y_{ij} at all levels and all matrixes is the product of the terms of Equation 22 over all $i = 1, \dots, k$ and $j = 1, \dots, q_i$,

$$\prod_{i=1}^k \prod_{j=1}^{q_i} f(y_{ij}; n_{ij}, d_{ij}, a_0, f_i) = \prod_{i=1}^k \prod_{j=1}^{q_i} \binom{n_{ij}}{y_{ij}} (1 - p_{ij})^{y_{ij}} p_{ij}^{n_{ij} - y_{ij}} \quad (24)$$

Given the sample size A_0 and the results of the experiment, (d_{ij}, n_{ij}, y_{ij}) ; $i = 1, \dots, k$; $j = 1, \dots, q_i$, Equation 24 is a function of the k unknown parameters f_i of the model, the Likelihood function

$$L(f_i; d_{ij}, n_{ij}, y_{ij}, a_0) = \prod_{i=1}^k \prod_{j=1}^{q_i} f(y_{ij}; n_{ij}, d_{ij}) \tag{25}$$

$$= \prod_{i=1}^k \prod_{j=1}^{q_i} \binom{n_{ij}}{y_{ij}} (1-p_{ij})^{y_{ij}} p_{ij}^{n_{ij}-y_{ij}}$$

The Maximum Likelihood estimates \hat{f}_i of the parameters f_i are the values f_i that maximize the Likelihood function of Equation 25. Because the extreme values of a function and the extreme values of the logarithm of that function have the same argument values, we use the logarithm of the Likelihood function, the Loglikelihood function, to find the location of the maxima:

$$l(f_i; d_{ij}, n_{ij}, y_{ij}, a_0) = \ln L(f_i; d_{ij}, n_{ij}, y_{ij}, a_0) = \sum_{i=1}^k \sum_{j=1}^{q_i} \ln f(y_{ij}; n_{ij}, d_{ij}, a_0, f_i) \tag{26}$$

$$= \sum_{i=1}^k \sum_{j=1}^{q_i} \left(\ln \binom{n_{ij}}{y_{ij}} + y_{ij} \ln p_{ij} - (n_{ij} - y_{ij}) \ln (1 - p_{ij}) \right)$$

The estimates \hat{f}_i are the values for which the partial derivatives of the Loglikelihood function with respect to f_i ,

$$\frac{\partial l}{\partial f_i} = \sum_{j=1}^{q_i} \left(\frac{y_{ij} \exp(-\exp(\eta_{ij})) \exp(\eta_{ij})}{1 - \exp(-\exp(\eta_{ij}))} - (n_{ij} - y_{ij}) \exp(\eta_{ij}) \right); i = 1, \dots, k \tag{27}$$

are 0, i.e.,

$$\sum_{j=1}^{q_i} \left(\frac{y_{ij} d_{ij}}{\exp(\exp(\hat{\eta}_{ij})) - 1} - (n_{ij} - y_{ij}) d_{ij} \right) = 0; i = 1, \dots, k \tag{28}$$

with $\exp(\exp(\hat{\eta}_{ij})) = \exp(\exp(a_0 + \hat{f}_i + \ln d_{ij})) = \exp(A_0 \hat{F}_i d_{ij})$.

Generally, this is a system of k equations for the k estimates $\hat{f}_i = \ln \hat{F}_i$; however, for our model, it simplifies to one equation for each \hat{f}_i . To find the root of a nonlinear equation of the form 28 is a simple numerical problem.

Asymptotically, i.e., for $n_{ij} \rightarrow \infty$, \hat{f}_i is normally distributed with expectation $E(\hat{f}_i) = f_i$ and SD (7).

$$\sigma_{\hat{f}_i} = \frac{1}{\sqrt{\sum_{j=1}^{q_i} E(t_{ij}^2)}} \tag{29}$$

where E is the symbol for the expectation and t_{ij}^2 is the j th term in Equation 27,

$$t_{ij} = A_0 F_i d_{ij} \left(\frac{y_{ij} (1 - p_{ij})}{p_{ij}} - (n_{ij} - y_{ij}) \right) \tag{30}$$

its square is

$$t_{ij}^2 = \left(A_0 F_i d_{ij} \right)^2 \left(\frac{y_{ij} (1 - p_{ij})}{p_{ij}} - (n_{ij} - y_{ij}) \right)^2 \tag{31}$$

$$= \frac{\left(A_0 F_i d_{ij} \right)^2}{p_{ij}^2} (y_{ij} - n_{ij} p_{ij})^2$$

Because y_{ij} is binomially distributed with $E(y_{ij}) = n_{ij} p_{ij}$ the expectation of $(y_{ij} - n_{ij} p_{ij})^2$ is

$$E\left((y_{ij} - n_{ij} p_{ij})^2 \right) = \text{Var}(y_{ij}) = n_{ij} p_{ij} (1 - p_{ij}) \tag{32}$$

and hence,

$$E(t_{ij}^2) = \frac{\left(A_0 F_i d_{ij} \right)^2}{p_{ij}^2} n_{ij} p_{ij} (1 - p_{ij}) \tag{33}$$

$$= \left(A_0 F_i d_{ij} \right)^2 n_{ij} \frac{1 - p_{ij}}{p_{ij}} = \frac{n_{ij} \left(A_0 F_i d_{ij} \right)^2}{\exp(A_0 F_i d_{ij}) - 1}$$

If we insert Equation 33 in Equation 29, we get

$$\sigma_{\hat{f}_i} = \frac{1}{\sqrt{\sum_{j=1}^{q_i} \frac{n_{ij} \left(A_0 F_i d_{ij} \right)^2}{\exp(A_0 F_i d_{ij}) - 1}}} \tag{34}$$

This asymptotic result is a very good approximation even for small values of the n_{ij} .