Electronic Supplementary Material (ESI) for Analytical Methods

A PROTOCOL FOR THE VALIDATION OF QUALITATIVE METHODS OF DETECTION: FURTHER EXAMPLES OF APPLICATION

Example 1: Detection of peanut protein

A collaborative trial (18 laboratories, 5 replicate samples at 7 levels including zero) of a dipstick to test to an allergen (peanut) in cookies[1] yielded the following results (Table 1). Estimates of statistical parameters and prediction intervals for the probability of a positive response are shown in Tables 2 and 3

Figure 1 shows that the observed probabilities of detection tell us that we can expect the limit of detection to be no higher than 29 mg/kg peanut in cookies for 95% of laboratories using this method, and limits of detection may be as low as 13 mg/kg in some laboratories. We can expect that the false positive probability will be less than 0.14 for 95% of laboratories Figure 6 shows that a large part of the observed variation can be explained by the uncertainty about the mean probability of detection and the random binomial sampling variation associated with 5 replicates per laboratory for most concentrations. Hence, if the uncertainty associated with the estimated limit of detection is too large then estimates of method performance may be improved by undertaking further studies using more replicate samples. This might be the case if the target limit of detection were around 20 mg/kg. The relatively large size of the random binomial sampling variation compared to the size of between laboratory variation for concentrations less than 5 mg/kg (Figure 2) may explain the reduction in the estimated upper limit for the probability of detection between 0 and 1.5 mg/kg.

Table 1: Number of positive responses out of 5 replicate tests for the detection of peanut protein in cookies

Lab	0 mg/kg	1.5 mg/kg	4 mg/kg	8.2 mg/kg	14 mg/kg	21 mg/kg	30 mg/kg
1	0	0	1	3	4	5	5
2	0	0	0	3	2	5	5
3	0	0	0	2	4	5	5
4	0	0	0	1	5	5	5
5	0	0	0	0	5	5	5
6	0	1	0	2	5	5	5
7	0	0	0	0	4	5	5
8	0	0	0	4	3	5	5
9	0	0	0	1	5	4	5
10	1	0	0	0	1	5	5
11	0	0	1	2	5	4	5
12	0	0	1	1	5	5	5
13	0	0	0	2	5	5	5
14	0	0	0	0	1	5	5
15	1	0	0	0	4	5	5
16	0	0	0	1	2	5	5
17	0	0	0	1	3	5	5
18	0	0	0	0	1	5	5

The effect of the low number of replicates is particularly clear for results produced by the analysis of cookies that did not contain peanut. Here the 2 out of 90 samples gave a positive response, giving an estimated average false positive probability of 0.022 with a 95% confidence interval of 0.005 to 0.069. The problem is that the estimated probability of a positive response for those two laboratories that each produced a single positive response is 0.2, and these high estimates have a big impact on the upper limit for the prediction interval for the probability of a positive response. In short the results are consistent with both a low-between-laboratory-variation false positive probability of less than 0.01 (probably good enough for a screening method), and a much higher (estimated upper 95% prediction interval of 0.14) false positive rate for some laboratories because the number of replicates per laboratory is low. In order to defend against these kinds of outcomes where preliminary work leads to the expectation of a few false positive or false negative results a larger number of replicate analyses should be undertaken in each laboratory (see Assessment of draft protocol performance and study design).

Table 2: Estimates of statistical parameters used to describe performance of a method to detect peanut protein in cookies

Concentration (mg/kg)	N	X	$\bar{p}$(1)	s_R	v_s(2)	w_s(3)	v_h(4)	w_h(5)
0	90	2	0.0222	0.0647	0.0932	4.101	2.5	88.5
1.5	90	1	0.0111	0.0471	0.0438	3.901	1.5	89.5
4	90	3	0.0333	0.0767	0.1493	4.329	3.5	87.5
8.2	90	23	0.2556	0.2455	0.5512	1.606	23.5	67.5
14	90	64	0.7111	0.3085	0.8240	0.3347	64.5	26.5
21	90	88	0.9778	0.0647	4.101	0.0932	88.5	2.5
30	90	90	1.0000	0.0000	NC	NC	90.5	0.5

N: Total number of analyses, X: total number of positive results; $\bar{p}$(1): estimated mean probability of detection (Equation 1); s_R: standard deviation of estimates of probability of detection from the individual laboratories; v_s(2): shape parameter of beta distribution based on mean and standard deviation of probability of detection (Equation 2); w_s(3): shape parameter of beta distribution based on mean and standard deviation of probability of detection (Equation 3); v_h(4): shape parameter of beta distribution based on N and X (Equation 4), w_h(5): shape parameter of beta distribution based on N and X (Equation 5). NC: Not calculated.

Table 3: Estimates lower (95%) and upper (5%) limits for the probability of detection in a laboratory using the method to detect peanut protein in cookies

Concentration (mg/kg)	L_s(6)	U_s(7)	L_h(8)	U_h(9)	Lowerlimit (10)	Upperlimit (11)
0	1.8E-15	0.1376	0.00640	0.0601	1.8E-15	0.0323
1.5	<7E-28	0.0622	0.00196	0.0426	<7E-28	0.5588
4	3.1E-10	0.1892	0.01214	0.0762	3.1E-10	0.9102
8.2	0.00252	0.7617	0.1863	0.3363	0.00252	0.9973
14	0.0848	0.9998	0.6283	0.7841	0.0848	1.0000
21	0.8624	1.0000	0.9398	0.9936	0.8624	1.0000
30	NC	NC	NC	NC	0.9673b	1.0000b

L_s(6): 5^{th} percentile of beta distribution with v_s and w_s shape parameters (Equation 6); U_s(7): 95^{th} percentile of beta distribution with v_s and w_s shape parameters (Equation 7); L_h(8): 5^{th} percentile of beta distribution with v_h and w_h shape parameters (Equation 8); U_h(9): 95^{th} percentile of beta distribution with v_h and w_h shape parameters (Equation 9); Lowerlimit(10): Lower limit of prediction interval for probability of detection when the method is applied in a laboratory (Equation 10); Upperlimit(10): Upper limit of prediction interval for probability of detection when the method is applied in a laboratory (Equation 11); b: these values were calculated using Equation 10b and 11b. NC: Not calculated

Figure 1: Estimates of an upper limit for the false positive probability and prediction interval for the limit of detection for peanut protein in cookies

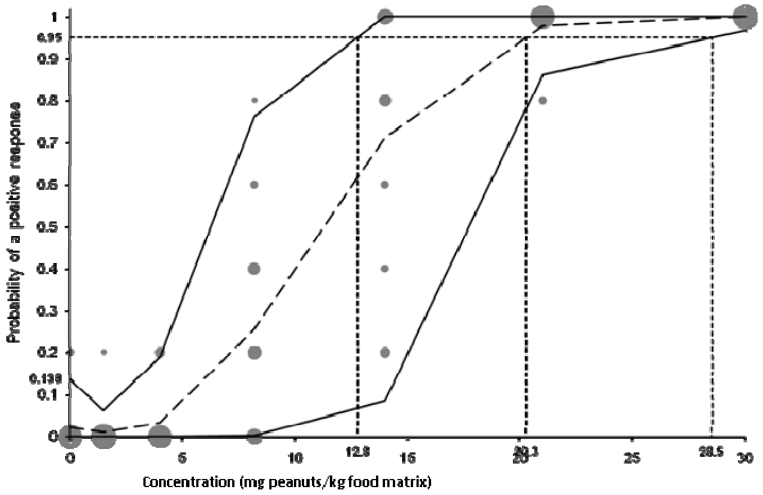
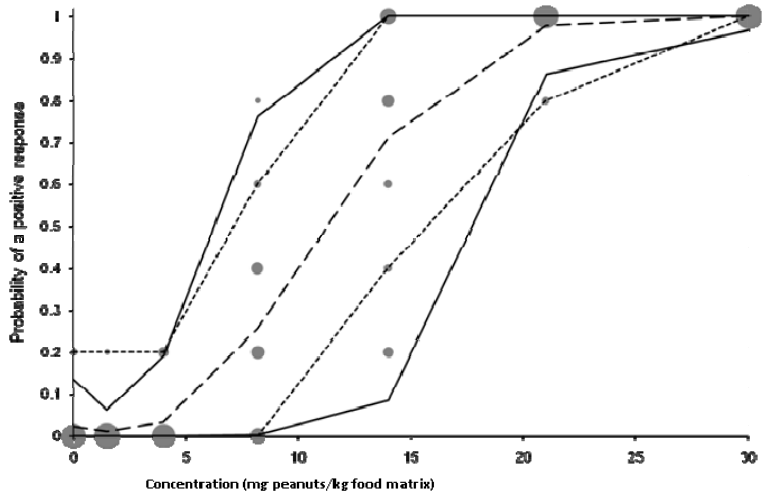


Figure 2: Comparison between the interval within which observed probabilities of detection lie and uncertainty associated with the use of a small number of replicate analyses



Probability of detection observed by 18 labs; by 9 labs, by 3 labs, by 1 lab (area proportional to number of labs), — — Mean probability of detection, ——— Lower (5%) and upper (95%) limits of a prediction interval for probability of detection in a lab. – – – lower(5%), expected and upper (95%) limits of a prediction interval for limit of detection in a lab; ..... Lower (5%) and upper (95%) for observed probabilities of detection if all observations are consistent with the estimated mean probability of detection with no between-laboratory variation

Example 2: A collaborative trial for the detection of salmonella in ground beef

A collaborative trial (11 laboratories, 6 replicate samples at 3 levels including zero) of a method for the detection of salmonella in ground beef yielded the following results[2] (Table 4), following the removal of results produced by one laboratory whose performance was judged to be inconsistent with the other laboratories.13 Calculation of the upper and lower limits (5 and 95%) for estimated probability of

detection at each concentration is shown in Tables 5 and 6. The results showed that a probability of detection was not demonstrated to be reliably (for at least 95% of laboratories) above 95% at any of the concentrations (Figure 3). We were able to estimate that the limit of detection, when the method is applied in a new laboratory, can be expected to be at least 8 cfu/25g in 95% of laboratories and that the false positive probability was likely to be less than 0.05.

The observed variation between laboratories can be explained by the uncertainty about the mean probability of detection and the binomial sampling variation associated with 6 replicates per laboratory for all concentrations (Figure 4).

Table 4: Number of positive responses out of 6 replicate tests for the presence of salmonella in ground beef

Lab	0 cfu/25 g	0.75 cfu/25 g	10.75 cfu/25 g
1	0	2	6
2	0	1	4
3	0	3	5
4	0	3	6
5	0	5	6
6	0	2	6
7	0	4	6
8	0	4	5
9	0	2	6
10	0	2	6

Table 5: Estimates of statistical parameters used to describe performance of a method to detect salmonella in ground beef

Concentration (cfu/25g)	*N*	*X*	$\bar{p}(1)$	s_R	$v_s(2)$	$w_s(3)$	$v_h(4)$	$w_h(5)$
0	60	0	0.0000	0.0000	NC	NC	0.5	60.5
0.75	60	28	0.4667	0.2049	2.300	2.63	28.5	32.5
10.75	60	56	0.9333	0.1165	3.343	0.2388	56.5	4.5

N: Total number of analyses, X: total number of positive results; $\bar{p}(1)$: estimated mean probability of detection (Equation 1); s_R: standard deviation of estimates of probability of detection from the individual laboratories; $v_s(2)$: shape parameter of beta distribution based on mean and standard deviation of probability of detection (Equation 2); $w_s(3)$: shape parameter of beta distribution based on mean and standard deviation of probability of detection (Equation 3); $v_h(4)$: shape parameter of beta distribution based on N and X (Equation 4), $w_h(5)$: shape parameter of beta distribution based on N and X (Equation 5).

Table 6: Estimates lower (95%) and upper (5%) limits for the probability of detection in a laboratory using the method to detect salmonella in ground beef

Concentration (cfu/25g)	$L_s(6)$	$U_s(7)$	$L_h(8)$	$U_h(9)$	*Lowerlimit (10)*	*Upperlimit (11)*
0	NC	NC	NC	NC	0[c]	0.0487[c]
0.75	0.1392	0.8107	0.3635	0.5722	0.1392	0.5588
10.75	0.6756	1.0000	0.8648	0.9719	0.6756	1.0000

$L_s(6)$: 5th percentile of beta distribution with v_s and w_s shape parameters (Equation 6); $U_s(6)$: 95th percentile of beta distribution with v_s and w_s shape parameters (Equation 7); $L_h(8)$: 5th percentile of beta distribution with v_h and w_h shape parameters (Equation 8); $U_h(9)$: 95th percentile of beta distribution with v_h and w_h shape parameters (Equation 9); Lowerlimit(10): Lower limit of prediction interval for probability of detection when the method is applied in a laboratory (Equation 10); Upperlimit(10): Upper limit of prediction interval for probability of detection when the method is applied in a laboratory (Equation 11); c: these values were calculated using Equation 10c and 11c

Figure 3: Estimates of an upper limit for the false positive probability and prediction interval for the limit of detection for salmonella in ground beef.

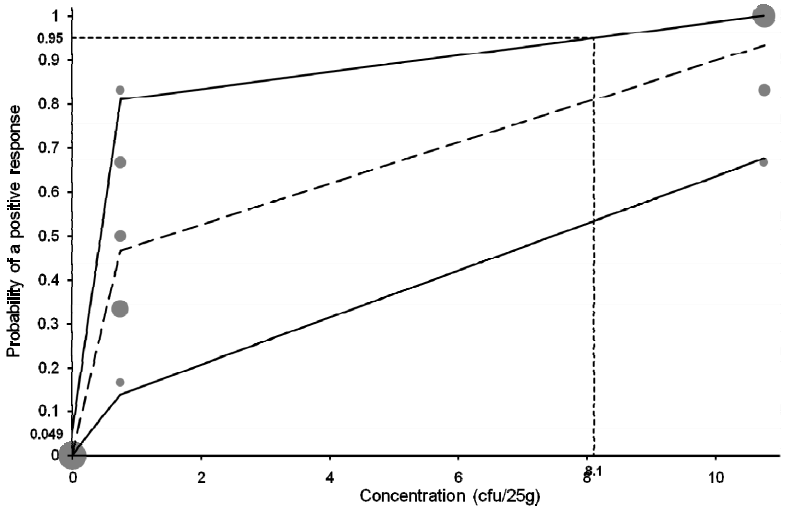
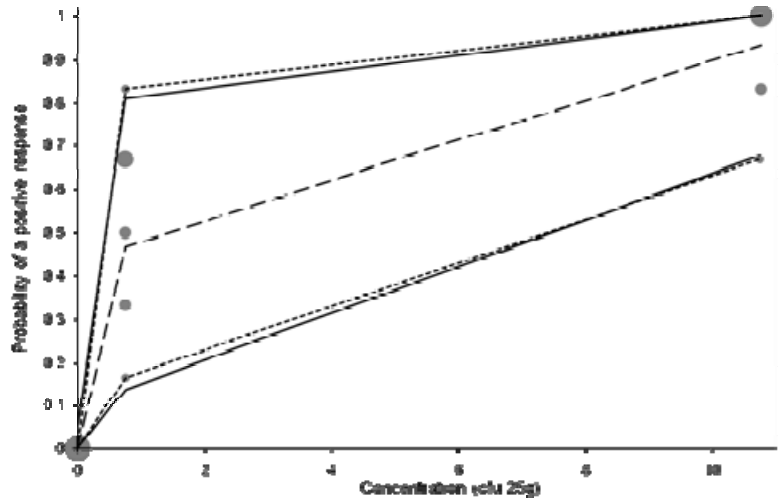


Figure 4: Comparison between the interval within which observed probabilities of detection lie and the uncertainty associated with the use of a small number of replicate analyses



Probability of detection observed by 11 labs; by 6 labs, by 2 labs, by 1 lab (area proportional to number of labs), — — Mean probability of detection, ——— Lower (5%) and upper (95%) limits of a prediction interval for probability of detection in a lab. – – – lower(5%), expected and upper (95%) limits of a prediction interval for limit of detection in a lab; ..... Lower (5%) and upper (95%) for observed probabilities of detection if all observations are consistent with the estimated mean probability of detection with no between-laboratory variation

Hence estimates of method performance may be improved by undertaking further studies using more replicate samples, possibly including some higher concentration samples. However, we can't be confident that this will be fruitful unless a limit of detection larger than 8 cfu/25g is considered fit for purpose.

References

1 A.J. van Hengel, C. Capelletti, M. Brohee, and E. Anklam, JAOAC Int., 2006, 89(2), 462-468

2 P. Wehling, R.A. Labudde, S.L. Brunelle, and M.T. Nelson, JAOAC Int., 2001, 94(1), 335-347