Epidemiol. Infect. (2014), **142**, 1061–1069. © Cambridge University Press 2013
doi:10.1017/S0950268813002173

Estimation of the sensitivity of environmental sampling for detection of *Salmonella* in commercial layer flocks post-introduction of national control programmes

M. E. ARNOLD[1]*, F. MARTELLI[2], I. McLAREN[2] AND R. H. DAVIES[2]

[1] *Animal Health and Veterinary Laboratories Agency (AHVLA), Sutton Bonington, Loughborough, UK*
[2] *Department of Bacteriology and Food Safety, AHVLA, Addlestone, Surrey, UK*

Received 23 May 2013; Final revision 23 July 2013; Accepted 14 August 2013; first published online 11 September 2013

SUMMARY

A key element of national control programmes (NCPs) for *Salmonella* in commercial laying flocks, introduced across the European Union, is the identification of infected flocks and holdings through statutory sampling. It is therefore important to know the sensitivity of the sampling methods, in order to design effective and efficient surveillance for *Salmonella*. However, improved *Salmonella* control in response to the NCP may have influenced key factors that determine the sensitivity of the sampling methods used to detect *Salmonella* in NCPs. Therefore the aim of this study was to compare estimates of the sensitivity of the sampling methods using data collected before and after the introduction of the NCP, using Bayesian methods. There was a large reduction in the sensitivity of dust in non-cage flocks between the pre-NCP studies (81% of samples positive in positive flocks) and post-NCP studies (10% of samples positive in positive flocks), leading to the conclusion that sampling dust is not recommended for detection of *Salmonella* in non-cage flocks. However, cage dust (43% of samples positive in positive flocks) was found to be more sensitive than cage faeces (29% of samples positive in positive flocks). To have a high probability of detection, several NCP-style samples need to be used. For confirmation of *Salmonella*, five NCP faecal samples for cage flocks, and three NCP faecal boot swab samples for non-cage flocks would be required to have the equivalent sensitivity of the EU baseline survey method, which was estimated to have an 87% and 75% sensitivity to detect *Salmonella* at a 5% within-flock prevalence in cage and non-cage flocks, respectively.

Key words: Animal pathogens, Bayesian methods, laying flocks, *Salmonella*, sampling.

INTRODUCTION

Salmonella Enteritidis (SE), and *S.* Typhimurium (ST) are responsible for the majority of cases of human salmonellosis in the UK and elsewhere in Europe [1]. Epidemiological investigations demonstrated that contaminated eggs produced by infected laying hens were the main source of human infection of SE [2, 3]; therefore it is important to reduce its incidence. In response, the national control programmes (NCPs) that have been introduced in EU Member States have been an important measure for reducing the *Salmonella* burden associated with food animal production. A critical element of these NCPs is the identification of infected poultry flocks and holdings so that sale of fresh eggs can be restricted and infection eliminated from the holding [4]. It is therefore important to know the sensitivity of sampling in

* Author for correspondence: Dr M. E. Arnold, Animal Health and Veterinary Laboratories Agency (AHVLA), The Elms, College Road, Sutton Bonington, Loughborough LE12 5RB, UK. (Email: mark.arnold@ahvla.gsi.gov.uk)

order to estimate the efficacy of detection and true prevalence in order to be able to design effective surveillance schemes.

Environmental sampling of poultry houses, in which contributions from a potentially large number of individual faecal droppings are pooled, is regarded as more cost-effective and sensitive than sampling faeces from a large number of individual birds [5]. The sampling adopted for the monitoring of *Salmonella* in commercial laying flocks in EU Member States consists of environmental sampling of faeces and dust from the flock, consisting of either naturally mixed faeces or boot swabs and dust sampled from different parts of the poultry house.

Studies have been performed to determine the sensitivity of various environmental sampling methods [6–10]. Included in these studies was the dust and faecal sampling methods, which have been used for NCP sampling by the competent authority, and the method used in the EU baseline survey [11] for *Salmonella* in commercial laying flocks, which is used in some Member States for confirmatory testing in order to exclude the possibility of a 'false-positive' initial test [12]. Studies have shown that sampling dust from flocks is a more sensitive method for detecting *Salmonella* than sampling faeces [6–10], and on this basis both the EU baseline survey method and the NCP method used by the competent authority include dust sampling, where dust is available.

A key feature of environmental sampling is that the sensitivity will depend upon the prevalence of infection in the flock being sampled [8, 9]. It is possible that the introduction of NCPs may have influenced the sensitivity of sampling. First, as farmers' efforts to control *Salmonella* could have reduced the within-flock prevalence, resulting in lower mean sensitivity of sampling methods. Second, other changes since the introduction of the NCP, such as the change to colony cages, may also have affected the sampling methods, so that the sensitivity for a given within-flock prevalence may also have changed. The objective of this study was to estimate the sensitivity of the NCP and EU survey sampling methods, for cage and non-cage flocks relative to the within-flock prevalence, and to determine whether there are differences in the sensitivity of the sampling methods since the introduction of the NCP, and thus to estimate the number of NCP samples required for optimal detection of *Salmonella* in cage and non-cage flocks. In addition, the sensitivity of an in-house method ('AHVLA sampling' [7]), which also involves environmental sampling of faeces and dust was also evaluated pre- and post-NCP.

METHODS

Environmental and individual bird sampling

Data to determine the sensitivity of individual and environmental sampling sensitivity arose from three studies: two of which were pre-NCP ([8], and R. H. Davies (AHVLA), unpublished data) and one of which was post-NCP [13]. All data from the pre-NCP studies were from flocks known to be infected with SE whereas all farms in the post-NCP study (20 flocks) [13] had been detected as *Salmonella* positive (not restricted to SE) through the NCP for *Salmonella* in laying hens. Environmental sampling in both studies consisted of three different methods: the method used in the EU baseline survey (the 'EU survey method') [11, 14], the method used for official sampling in the NCP for *Salmonella* in laying flocks (the 'NCP sampling method' [15] and an in-house method ('AHVLA sampling') [7]. Briefly, the EU survey method involved seven tests, consisting of five 200–300 g composite faecal samples (for caged flocks) or five pairs of boot swabs (for free-range or barn flocks), each representative of one fifth of the laying house and 2 × 250 ml dust samples. The NCP sampling method in non-cage flocks consisted of two tests, one faecal sample formed from 2 × 150 g faecal samples, each representing half of the house and collected from the same locations as the EU survey method, and one sample of 250 ml (100 g) dust, collected from prolific sources of dust throughout the house. In non-cage flocks, two pairs of boot swabs were used in place of the 2 × 150 g faecal samples, each pair collected from a representative half of the house. The AHVLA sampling method consisted of 10–20 composite faecal samples, each weighing ~25 g, and 10–20 dust samples, each weighing ~15 g, from representative point locations across the house, collected with buffered peptone water (BPW)-impregnated gauze swabs (Robinson Healthcare, UK). For the individual bird sampling in both studies, hens were culled by cervical dislocation on farm, and stored at 4 °C overnight. At post-mortem examination, ovaries/oviduct and caeca were aseptically removed and cultured separately as 25 g samples in 225 ml BPW.

Furthermore, for the post-NCP study [13] 100 individual freshly voided droppings were tested both individually and in pools of five, for which a

semi-quantitative dilution-enrichment method was used. Briefly, 100 ml of the initial BPW solution of 10 g faecal sample was used to make a tenfold dilution series in BPW to 10^{-7}. The last dilution to test positive was recorded [16]. For the earlier study, 60 pools of five were tested via the semi-quantitative dilution-enrichment method but individual droppings were not tested. The unpublished data consisted of EU sampling (five faecal, two dust), AHVLA sampling (20 faecal, 20 dust), and NCP sampling (one faecal, one dust) in seven cage flocks and one non-cage flock.

Statistical models to estimate sensitivity of sampling

To perform comparisons between the sensitivity of sampling pre- and post-NCP, it is important to be able to account for differences in within-flock prevalence between the studies, so that it can be determined whether the observed differences are merely due to changes in the within-flock prevalence. These environmental sampling data allow the sensitivity of each sampling method relative to the within-flock prevalence to be determined. The per-sample sensitivity of the dust and faecal sampling for both the EU survey method and the NCP official sampling was determined by Bayesian methods [8, 9] using the data on the environmental sampling from each farm.

Model for the individual test sensitivity of caeca and ovaries/oviduct

Denoting the sensitivity of caeca and ovaries/oviduct by Se_1, Se_2, respectively, the likelihood of the results for each bird is given by:

$$P_{00} = \pi(1 - Se_1)(1 - Se_2) + \theta,$$

$$P_{10} = \pi Se_1(1 - Se_2) - \theta,$$

$$P_{01} = \pi(1 - Se_1)Se_2 - \theta,$$

$$P_{11} = \pi Se_1 Se_2 + \theta,$$

where P_{00} represents the likelihood that the tests on both caeca and ovaries/oviduct are negative, P_{10} the likelihood that the test on caeca is positive and the test on the ovaries/oviduct is negative, θ is the covariance between the tests of ovaries/oviduct and caeca, and π represents the prevalence of *Salmonella* infection (within the flock). The likelihood of the data for the ovaries/oviduct and caeca results for each flock then follows a multinomial distribution, with the cell probabilities given by the Pij values above and n=the number of birds tested in the flock.

Model for the composite sampling methods

We follow the model developed in [8] which allows for the possibility that the probability of an environmental faecal or dust sample testing positive is dependent on the prevalence of *Salmonella* infection in the house from which it was sampled, by assuming that each follow a logistic regression curve, i.e. denoting the sensitivity of method i by η_i

$$\eta_i(\pi) = \frac{\exp(\alpha_i + \beta_i \pi)}{1 + \exp(\alpha_i + \beta_i \pi)},$$

where π is as defined earlier and α_i and β_i are the unknown parameters of the logistic regression model, with the parameter β_i representing the dependence of the method on the within-flock prevalence. There were six types of sample for the three detection methods: each of the EU, NCP and AHVLA methods having both a dust and faecal sample, so there were six values of α, β to be estimated. In addition, the sensitivity of testing individual faecal droppings was shown to be dependent on flock-level prevalence [9] and so this was also assumed to follow a logistic regression model.

The specificity of the culture methods was assumed to be 100%. The number of positives for farm j for each environmental sample test i followed a binomial distribution with $p = \eta_i(\pi_j)$ and n=number of samples for sample type i on farm j.

The estimation of the unknown parameters Se_1, Se_2, π_j, α_i, β_i was performed using WinBUGS v. 3.1. The parameters Se_1, Se_2, and π_j were each assumed to follow a beta distribution, which is a flexible distribution and is constrained to take values between 0 and 1, and is therefore ideal at representing test sensitivity and prevalence. Each α_i, β_i was assumed to be normally distributed, with mean 0 and variance 1000, i.e. non-informative priors were used throughout.

It was considered important to estimate the sensitivity of sampling separately for cage and non-cage flocks, as important differences in the sensitivity found between cage and non-cage flocks have previously been reported [7]. Therefore hypothesis tests were performed to determine whether there was any significant difference between the parameters determining test sensitivity between (i) the cage ($n = 32$) and non-cage flocks ($n = 17$) as differences had been found in a previous study [7] and (ii) between sampling carried out pre-NCP ([8], and R. H. Davies (AHVLA), unpublished data) and post-NCP [13], as the introduction of the NCP may have influenced

the sensitivity of the sampling methods. These hypothesis tests were performed by use of the Deviance Information Criterion (DIC) [17], which is a Bayesian analogue of Akaike's Information Criterion. To assist in the interpretation of the DIC, a DIC weight (w_{DIC}) was calculated for each model being compared, which gives an estimate of the probability that each model is the best model for the data at hand, and is calculated according to

$$w_{\mathrm{DIC}} = \frac{\exp(-\frac{1}{2}\Delta \mathrm{DIC})}{\sum \exp(-\frac{1}{2}\Delta \mathrm{DIC})},$$

where ΔDIC was the difference between the model in question and the minimum value of the DIC for the models being compared, and the denominator was the sum of the differences over all the models being compared. The best-fitting model out of those compared will be that with the highest DIC weight, and a value close to 1 indicates strong evidence that it is the best model.

All calculations were performed in WinBUGS v. 3.1 [18], using a burn-in of 5000 iterations followed by 10 000 iterations of the model. Inspection of the history of each parameter and the use of the Gelman–Rubin statistic [19] were used to check convergence.

Number of samples required for detection of *Salmonella*

The sensitivity of the EU survey method was then calculated from the per-sample estimates of the dust and faecal sampling (i.e. the probability of detecting *Salmonella* in a flock after taking two EU dust and five EU faecal samples). The sensitivity of the NCP sampling method was then calculated for various numbers of samples for the following scenarios:

(1) Assuming that two NCP dust samples are included, and increasing the number of NCP faecal samples.
(2) Assuming that only NCP faecal samples are taken.

The minimum number of samples for which the sensitivity of the NCP sampling method was greater than the EU survey method was also calculated.

Estimation of c.f.u./g in faecal samples

The c.f.u./g was estimated in individual samples from enumeration of the pools of five using the maximum-likelihood approach developed in a previous study [8]. This method assumes that the $\log_{10}$ c.f.u./g follows a normal distribution (i.e. the c.f.u./g follows a lognormal distribution). The key element of the method is the approximating assumption that the $\log_{10}$ count gives the count of the most contaminated sample in the pool, with other counts being lower. The method uses the estimated within-flock prevalence from the estimation in WinBUGS to infer the probability distribution of the number of positive samples in each pool, and accounts for this in the estimation. A likelihood-ratio test was performed to detect whether there were any differences in the parameters of the lognormal distribution underlying the *Salmonella* count in faecal samples between the pre- and post-NCP studies.

RESULTS

Comparison of sensitivity on a per-sample basis in cage *vs.* non-cage flocks

When looking at the total across both the pre- and post-NCP studies on a per-sample basis, a higher proportion of faecal (bootswab) samples were positive in non-cage compared to cage flocks (see Table 1 for summary; flock-level details from pre-NCP studies ([8] and R. H. Davies (AHVLA), unpublished data); and the post-NCP study [13] in Supplementary Tables S1–S3), for all sample protocols (i.e. NCP, EU and AHVLA sampling). In contrast, a higher proportion of dust samples were positive in cage than in non-cage flocks for all sample protocols. For faeces, NCP sampling had the highest proportion of samples that tested positive, whereas for dust EU sampling had the highest proportion of positive samples, with individual AHVLA dust samples having the lowest proportion of positive samples.

When comparing the proportion of samples positive between the pre- and post-NCP studies, the post-NCP studies generally had a lower proportion of samples positive compared to the pre-NCP studies. However, there was a particularly large reduction in the proportion of non-cage dust samples positive in the post-NCP study compared to the pre-NCP studies.

For the most part, the model estimates of sensitivity *vs.* within-flock prevalence (Fig. 1) showed a consistent order in the relative sensitivity of sampling dust/faeces in cage/non-cage flocks for NCP, EU and AHVLA sampling: cage dust was more sensitive than cage faeces, and dust from non-cage flocks was the least sensitive for a given within-flock prevalence for NCP, EU and AHVLA sampling. The exception

Table 1. *The total number of samples positive for dust and faeces from NCP, EU and AHVLA sampling for* Salmonella *in cage and non-cage flocks, across a total of 41 flocks*

		Faecal samples		Dust samples	
Sample type	Period	Cage	Non-cage	Cage	Non-cage
NCP	Pre-NCP	11/17 (64·7%)	6/6 (100%)	10/17 (58·8%)	6/6 (100%)
	Post-NCP	14/46 (30·4%)	40/60 (66·7%)	20/46 (43·5%)	5/30 (16·7%)
	Total	25/63 (39·7%)	46/66 (69·7%)	31/63 (49·2%)	11/36 (30·6%)
EU	Pre-NCP	38/105 (36·2%)	27/40 (67·5%)	25/42 (59·5%)	12/16 (75%)
	Post-NCP	14/55 (25·5%)	23/45 (51·1%)	10/22 (45·5%)	2/18 (11·1%)
	Total	52/160 (32·5%)	50/85 (58·8%)	35/64 (54·7%)	14/34 (41·2%)
AHVLA	Pre-NCP	103/370 (27·8%)	74/190 (38·9%)	194/390 (49·7%)	63/90 (70%)
	Post-NCP	48/200 (24%)	79/180 (43·9%)	46/200 (23%)	14/180 (7·8%)
	Total	151/570 (26·5%)	153/370 (41·4%)	240/590 (40·7%)	77/270 (28·5%)

NCP, National control programme; EU, European Union; AHVLA, Animal Health and Veterinary Laboratories Agency.

to this consistent ordering in terms of sensitivity was faeces in non-cage flocks, which was the most sensitive method for AHVLA sampling, but not for EU and NCP sampling (note that, while a higher proportion of non-cage faeces were positive overall compared to cage dust, the lower within-flock prevalence in cage flocks means that the cage dust is more sensitive than non-cage faeces, once within-flock prevalence has been accounted for).

Use of the DIC indicated a significant difference between the sensitivity of the tests for cage *vs.* non-cage flocks ($P < 0·001$). There was also strong evidence of differences in the sensitivity of sampling between the pre- and post-NCP studies ($P < 0·001$), with the best-fitting model being one where dust sampling (all sample protocols, both cage and non-cage) and sampling of non-cage faeces (all sample protocols) differed between pre- and post-NCP studies. Therefore, sensitivities were estimated separately for cage and non-cage flocks, using only data from the post-NCP study [13] to estimate the sensitivity of dust sampling and non-cage faeces, for the purposes of sample size estimation. The best-fitting model also had common sensitivity of sampling individual droppings between pre- and post-NCP studies.

In both the pre- and post-NCP studies there was a much higher within-flock prevalence of infected birds within non-cage flocks than in cage flocks, 35·1% (non-cage) compared to 15·9% (cage) in the pre-NCP study [8], and 15·5% compared to 5·5% in the post-NCP study [13].

The sensitivity of testing caeca was estimated to be 75·8% [95% credible interval (CrI 72·3–78·3)] across both pre- and post-NCP studies, and the sensitivity of testing ovary/oviduct samples was 67·2% (95% CrI 64·6–69·9).

Number of samples required for confirmatory sampling

For a given within-flock prevalence, fewer NCP samples are required to detect *Salmonella* in non-cage flocks compared to cage flocks (Table 2), because of the slightly higher estimated sensitivity of faeces in non-cage flocks compared to cage flocks. There was limited benefit in taking additional dust samples; while in cage flocks dust had higher sensitivity than faecal samples (Fig. 1), the increase in sensitivity of dust *vs.* faeces was not sufficient to reduce the overall number of samples to detect *Salmonella*, except at relatively high within-flock prevalence (Table 2). For non-cage flocks, taking two NCP dust samples had little or no impact on the number of NCP faecal samples required to detect a positive flock, reducing it by at most one, and often not reducing the number required at all (Table 2).

The sensitivity of the EU survey method to detect positive cage and non-cage flocks (i.e. at least one positive sample out of five EU faecal and two EU dust samples) at a range of within-flock prevalence is given in Table 3. To match the EU survey method for detecting *Salmonella*-infected farms, five NCP faecal samples or three NCP faecal and two NCP dust samples for cage flocks would be required in cage flocks. To match the EU survey method in non-cage flocks, three NCP faecal (boot swab) samples or two NCP faecal and two NCP dust samples would be

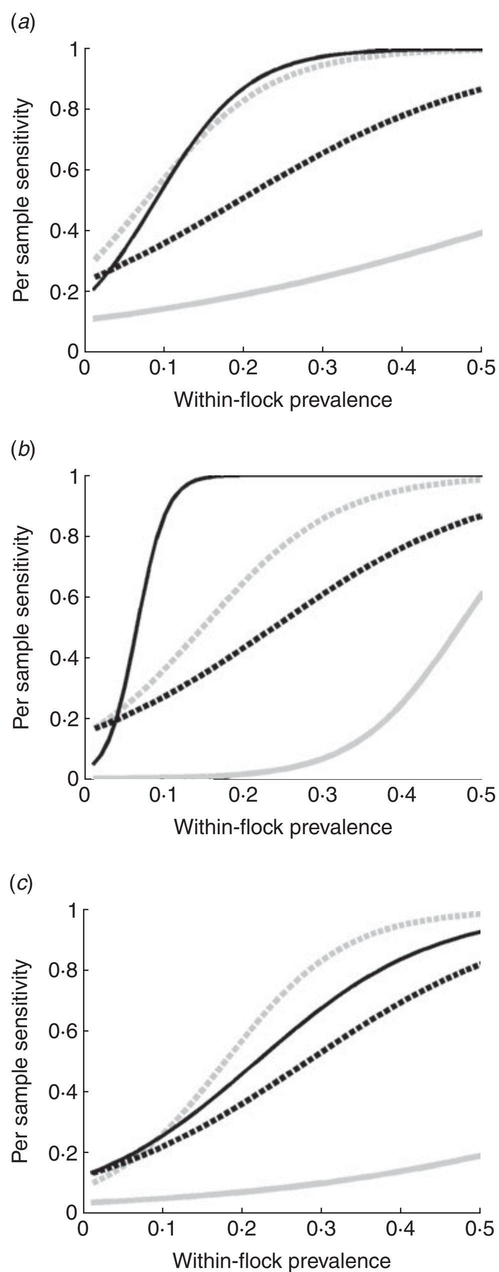

Fig. 1. The per-sample sensitivity of dust in cage (solid black lines) and non-cage (solid grey lines) flocks, and faeces in cage (dotted black lines) and non-cage (dotted grey lines) flocks relative to the within-flock prevalence for (*a*) NCP sampling; (*b*) EU sampling; (*c*) AHVLA sampling.

required (note that the EU survey method has higher sensitivity in cage flocks for a given within-flock prevalence due to the higher sensitivity of sampling dust in cage flocks compared to non-cage flocks, so the target sensitivity is higher in cage flocks to match the EU survey method).

Table 2. *Number of national control programme (NCP) faecal samples required in order to detect* Salmonella *with a 95% probability for cage and non-cage flocks, according to whether none or two NCP dust samples are taken in parallel*

Cage/ non-cage	No. of dust samples	Within-flock prevalence (%)					
		0·1	0·5	1·0	2·0	5·0	10·0
Cage	0	12	12	11	11	9	7
Cage	2	10	10	10	9	7	4
Non-cage	0	10	10	9	8	6	4
Non-cage	2	9	9	8	8	6	4

Table 3. *The estimated sensitivity of the European Union method to detect* Salmonella-*infected flocks, for cage and non-cage production types*

Cage/ non-cage	Within-flock prevalence (%)					
	0·1	0·5	1·0	2·0	5·0	10·0
Cage	0·62	0·64	0·65	0·71	0·87	0·98
Non-cage	0·59	0·61	0·62	0·65	0·73	0·84

Estimation of c.f.u./g in faeces

Count data from previous studies are given in Supplementary Table S4 for the pre-NCP study [8] and from [13] (Supplementary Table S5) for the post-NCP study [13]. While there was a higher proportion of pools of five faecal samples with higher counts in the pre-NCP study, when the within-flock prevalence was accounted for, there was no significant difference in the estimated c.f.u./g in individual samples either between the two studies ($P = 0{\cdot}65$) or between SE and non-SE ($P = 0{\cdot}57$), consistent with the finding of a common sensitivity of sampling individual droppings between the pre- and post-NCP studies. There was a good fit of the model to the observed counts of *Salmonella* in the pools of five (Fig. 2*a*). The mean c.f.u./g in individual samples was estimated to have a very low mean with very high variance. This indicates that the majority of samples had very low or zero counts of *Salmonella* but with a low proportion of samples having very high counts (Fig. 2*b*).

DISCUSSION

This study has shown important differences in the sensitivity of sampling between cage and non-cage flocks,

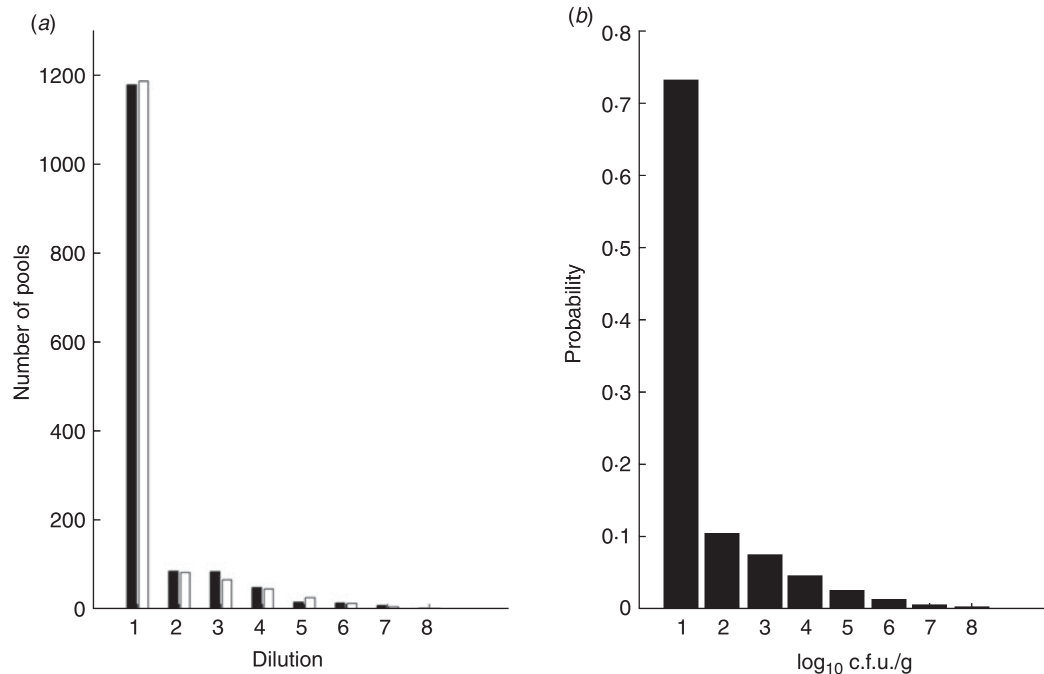

Fig. 2. (*a*) Comparison between the model and the observed number of pools of five positive for *Salmonella* at each tenfold dilution from 41 flocks sampled; and (*b*) the estimated c.f.u./g of *Salmonella* in individual samples.

for a given within-flock prevalence. While other studies have estimated the sensitivity of environmental sampling in cage and non-cage flocks [7, 10, 20], the data used in this study have individual bird data collected in parallel with the environmental sampling data, enabling a simultaneous estimate of the within-flock prevalence. This means that the impact of within-flock prevalence on the sensitivity of sampling can be accounted for, which is important when comparing the results of studies where prevalence is very different between them. By making use of data from more than one study, there was sufficient power to be able to divide the flocks into cage and non-cage, which was not possible with each study in isolation without greatly compromising statistical power.

There was a large reduction in within-flock prevalence between pre-NCP flocks and post-NCP flocks. The flocks in these studies were known positives, and therefore may represent a biased sample, since flocks with a prevalence too low to be detected would thus not be represented, and therefore the within-flock prevalence found in each study is possibly an overestimate of that in the population of *Salmonella*-infected flocks. However, it is likely that improved pest control, better vaccination of flocks and more regular dust removal in poultry houses since the introduction of the NCP will have resulted in a lower average within-flock prevalence in infected flocks [21, 22], and thus the lower within-flock prevalence in the post-NCP study is likely to be a reflection of lower within-flock prevalence in the population, compared to pre-NCP. This is important in terms of sampling, because it means that the sensitivity of environmental sampling methods will be lower than that reported in studies conducted prior to the introduction of the NCP [7, 10]. This could be one reason why the proportion of samples positive was higher for the pre-NCP studies compared to the post-NCP studies (Table 1).

Previous studies have shown that dust is more sensitive than faeces [7, 8]. However in the present study this was only true for cage flocks; for non-cage flocks faeces were more sensitive than dust for all three environmental sampling methods. The likely reason for this difference is the more frequent dust removal carried out in poultry houses since the introduction of the NCP, plus improved control of mice whose contaminated faeces and accumulated urine pillars may contribute to the occurrence of *Salmonella*-positive dust in a house where the birds are infected [16, 23–26].

Previous studies have found that faecal sampling in cage flocks is more sensitive than faecal sampling in non-cage flocks [7] in that there is a greater likelihood

of detecting positive cage flocks compared to non-cage flocks. In the present study, the opposite was found, as a higher proportion of faecal samples were positive in non-cage flocks compared to cage flocks in both the pre- and post-NCP studies (Table 2). This is likely to be influenced by the replacement of difficult to sample California style A-frame cage houses with colony cage houses in anticipation of the ban on conventional cages from the beginning of 2012 and inclusion of flocks with low levels of environmental contamination in earlier studies, in which the evaluation of detection of infection in 4000 pooled eggs from lightly contaminated laying houses was the main objective of the study.

The within-flock prevalence in non-cage flocks was much higher than that in cage flocks for both the pre- and post-NCP studies. This is likely to be due to a greater contact rate between infected birds in non-cage flocks compared to cage flocks and the inability to separate birds from their faeces [27]. This has important implications for the sensitivity of sampling. While there is uncertainty in the within-flock prevalence of *Salmonella vs.* the age of the flock, making the estimation of sample sizes based on a specific target prevalence problematical, the lower *Salmonella* prevalence in cage flocks will mean that a greater number of samples will be required for a high probability of detection. In particular, confirmation by sampling ova/oviduct and caeca or sampling of eggs is likely to be more effective in non-cage flocks than cage flocks, as the effectiveness of these methods will be directly proportional to the within-flock prevalence and level of environmental contamination. As for confirmation using NCP sampling methods, the present study (Table 2) suggests that several NCP samples would be required to have a high probability of detection. In the context of the NCP in each Member State, where each eligible commercial layer flock could be sampled up to four times per annum, this is likely to be impractical on each occasion, but theoretically, regular sampling should have a cumulative effect in detecting positive flocks. Even if just performed for official NCP sampling, where one flock is sampled per annum on holdings with >1000 birds, additional sampling would represent a large additional cost. However, it is valuable to apply additional sampling for confirmatory testing of flocks which were detected as positive through operator NCP sampling or for suspect flocks. The difference in the sensitivity of sampling of dust between cage and non-cage flocks is important to account for in such sampling, since the present study predicts that dust will be relatively ineffective for confirming *Salmonella* in non-cage flocks; in non-cage flocks it is therefore recommended that only faeces are sampled to optimize the likelihood of successful detection of *Salmonella.* The per-sample superiority of NCP samples compared to EU survey samples is likely to relate to the wider area of the house over which each NCP sample is taken and the direct mixing of the sample rather than dilution prior to mixing.

This study has shown important differences in the sensitivity of environmental sampling methods between cage and non-cage flocks, resulting in different numbers of samples required to have a >95% probability of detecting *Salmonella*. In particular, the sensitivity of dust sampling is very low in non-cage flocks and replacing dust samples with faecal samples is likely to improve the probability of detection of *Salmonella* in these flocks. For confirmation of *Salmonella*, taking several NCP-style samples would enable NCP sampling to have the equivalent power of detection of the EU survey method. Five NCP faecal samples for cage flocks, and three NCP faecal samples (i.e. three pairs of boot swabs) for non-cage flocks are required for this.

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0950268813002173.

ACKNOWLEDGEMENTS

This work was funded by the UK Department for Environment, Food and Rural Affairs projects OZ0332 and FZ2000.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Anon.** The Community Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents, Antimicrobial Resistance and Foodborne Outbreaks in the European Union in 2006. *EFSA Journal* 2007; **130**: chapter 3·1 pp. 24–105.
2. **Coyle EF, *et al.*** *Salmonella enteritidis* phage type 4 infection: association with hen's eggs. *Lancet* 1988; **2**: 1295–1297.
3. **Gillespie IA, *et al.*** Foodborne general outbreaks of *Salmonella* Enteritidis phage type 4 infection, England

and Wales 1992–2002: where are the risks? *Epidemiology and Infection* 2005; **133**: 795–801.

4. **Anon.** Decision (EC) No 2160/2003 of 17 November 2003 on the control of Salmonella and other specified food-borne zoonotic agents.
5. **Aho M.** Problems of *Salmonella* sampling. *International Journal of Food Microbiology* 1992; **15**: 225–235.
6. **Davies RH, Breslin M.** Environmental contamination and detection of Salmonella enterica serovar enteritidis in laying flocks. *Veterinary Record* 2001; **149**: 699–704.
7. **Carrique-Mas JJ, *et al.*** Comparison of environmental sampling methods for detection of salmonella in commerical flocks in the UK. *Letters in Applied Microbiology* 2008; **47**: 514–519.
8. **Arnold ME, Carrique-Mas JJ, Davies RH.** Estimation of the sensitivity of environmental sampling for detection of *Salmonella* Enteriditis in commercial egg-laying flocks relative to the within-flock prevalence. *Epidemiology and Infection* 2010; **138**: 330–339.
9. **Arnold ME, *et al.*** A comparison of pooled and individual bird sampling for detection of salmonella in commercial egg laying flocks. *Preventive Veterinary Medicine* 2011; **99**: 176–184.
10. **Mahe A, *et al.*** Bayesian estimation of flock-level sensitivity of detection of *Salmonella* spp., Enteritidis and Typhimurium according to the sampling procedure in French laying-hen houses. *Preventive Veterinary Medicine* 2008; **84**: 11–16.
11. **EFSA.** Report of the Task Force on Zoonoses Data Collection on the analysis of the baseline study on the prevalence of *Salmonella* in holdings of laying hen flocks of *Gallus gallus*. *EFSA Journal* 2007; **97**: 1–84.
12. **Anon.** Commission Regulation (EC) No. 1237/2007 of 23 October 2007 amending Regulation (EC) No. 2160/2003 of the European Parliament and of the Council and Decision 2006/696/EC as regards the placing on the market of eggs from *Salmonella* infected flocks of laying hens.
13. **Arnold ME, *et al.*** Estimation of the rate of egg contamination from *Salmonella* infected chickens. *Zoonoses and Public Health* (in press).
14. **EFSA.** Preliminary Report on the Analysis of the baseline study on the prevalence of *Salmonella* in laying hen flocks of *Gallus gallus*. *EFSA Journal* 2006; **81**: 1–71.
15. **Department for Environment, Food and Rural Affairs (Defra).** UK national control programme for *Salmonella* in layers (*Gallus gallus*), Defra, 2007.
16. **Wales A, Breslin M, Davies R.** Semiquantitative assessment of the distribution of *Salmonella* in the environment of caged layer flocks. *Journal of Applied Microbiology* 2006; **101**: 309–318.
17. **Spiegelhalter DJ, *et al.*** Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* 2002; **64**: 583–640.
18. **Lunn DJ, *et al.*** WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 2000; **10**: 325–337
19. **Brooks SP, Gelman A.** Alternative methods for monitoring convergence of iterative simulations. *Journal of Compuational and Graphical Statistics*. 1998; **7**: 434–455.
20. **Van Hoorebeke S, *et al.*** Faecal sampling underestimates the actual prevalence of *Salmonella* in laying hen flocks. *Zoonoses and public health* 2009; **56**: 471–476.
21. **Van Hoorebeke S, *et al.*** Determination of the within and between flock prevalence and identification of risk factors for *Salmonella* infections in laying hen flocks housed in conventional and alternative systems. *Preventive Veterinary Medicine* 2010; **94**: 94–100.
22. **Carrique-Mas JJ, *et al.*** Persistence and clearance of different *Salmonella* serovars in buildings housing laying hens. *Epidemiology and Infection* 2008; **19**: 1–10.
23. **Garber L, *et al.*** *Salmonella enterica* serotype *enteritidis* in table egg layer house environments and in mice in US layer houses and associated risk factors. *Avian Diseases* 2003; **47**: 134–142.
24. **Davies R, Breslin M.** Environmental contamination and detection of *Salmonella enterica* serovar *enteritidis* in laying flocks. *Veterinary Record* 2001; **149**: 699–704.
25. **Davies RH, Wray C.** Mice as carriers of *Salmonella enteritidis* on persistently infected poultry units. *Veterinary Record* 1995; **137**: 337–341.
26. **Henzler D J, Opitz HM.** The role of mice in the epizootiology of *Salmonella enteritidis* infection on chicken layer farms. *Avian Diseases* 1992; **36**: 625–631.
27. **Holt PS, *et al.*** The impact of different housing systems on egg safety and quality. *Poultry Science* 2011; **90**: 251–262.