Seroincidence of non-typhoid *Salmonella* infections: convenience *vs.* random community-based sampling

H.-D. EMBORG[1]*, J. SIMONSEN[2], C. S. JØRGENSEN[3], L. H. HARRITSHØJ[4],
K. A. KROGFELT[5], A. LINNEBERG[6,7,8] AND K. MØLBAK[1]

[1] *Department of Infectious Disease Epidemiology, Statens Serum Institut, Denmark*
[2] *Department of Epidemiology Research, Statens Serum Institut, Denmark*
[3] *Department of Microbiological Diagnostics and Virology, Statens Serum Institut, Denmark*
[4] *Department of Clinical Immunology, Capital Region, Copenhagen University Hospital, Denmark*
[5] *Department of Microbiology and Infection Control, Statens Serum Institut, Denmark*
[6] *Research Centre for Prevention and Health, Glostrup University Hospital, Denmark*
[7] *Department of Clinical Experimental Research, Glostrup University Hospital, Denmark*
[8] *Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark*

SUMMARY

The incidence of reported infections of non-typhoid *Salmonella* is affected by biases inherent to passive laboratory surveillance, whereas analysis of blood sera may provide a less biased alternative to estimate the force of *Salmonella* transmission in humans. We developed a mathematical model that enabled a back-calculation of the annual seroincidence of *Salmonella* based on measurements of specific antibodies. The aim of the present study was to determine the seroincidence in two convenience samples from 2012 (Danish blood donors, $n = 500$, and pregnant women, $n = 637$) and a community-based sample of healthy individuals from 2006 to 2007 ($n = 1780$). The lowest antibody levels were measured in the samples from the community cohort and the highest in pregnant women. The annual *Salmonella* seroincidences were 319 infections/1000 pregnant women [90% credibility interval (CrI) 210–441], 182/1000 in blood donors (90% CrI 85–298) and 77/1000 in the community cohort (90% CrI 45–114). Although the differences between study populations decreased when accounting for different age distributions the estimates depend on the study population. It is important to be aware of this issue and define a certain population under surveillance in order to obtain consistent results in an application of serological measures for public health purposes.

Key words: *Salmonella*, serology, surveillance.

INTRODUCTION

In 2011, salmonellosis was the second most common reported foodborne zoonotic disease in humans in the European Union with 95 548 cases reported to ECDC [1]. These figures underestimates the real incidence of infections as the number of reported cases depends on healthcare-seeking behaviour as well as the organization of the healthcare systems including laboratory services and practices in the different countries. *Salmonella* infections in humans are most often diagnosed by bacterial culture of faecal samples which requires the shedding of viable bacteria which

* Author for correspondence: Dr H.-D. Emborg, Department of Infectious Disease Epidemiology, Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen, Denmark.
(Email: hde@ssi.dk)

may decrease depending on the time since onset of infection. By contrast specific antibodies against a pathogen will be present in serum for at least some months, depending on the antibody isotype. We have previously shown that IgM and IgA antibodies against *Salmonella* persist for 2–4 months, whereas IgG antibodies may persist for up to 12 months [2]. On this basis it has been suggested that seroepidemiological methods can be applied as an active surveillance tool to measure the force of *Salmonella* transmission at the population level [3].

A mathematical model was previously developed to estimate *Salmonella* seroincidence in humans. In principle, the time since last infection recognized by the immune system (and the associated variability) can be calculated based on a set of IgG, IgM and IgA values measured in a single serum sample [4]. Using this model, the individual estimates of time since infection in each serum sample in a cross-sectional collection is converted to an annual seroincidence estimate in the population [5, 6]. This method can be applied to compare the force of infection independent of healthcare systems and laboratory surveillance in different countries.

There are several challenges that must be dealt with in order to make this principle feasible. Our previous work was to a large extent based on the analysis of collections of serum samples that were representative of general healthy populations [5, 6]. However, in many countries, such collections are not available or are reserved for other scientific purposes. Hence, in order to determine whether seroepidemiology is a feasible and practical tool, we investigated whether the analysis of sera from convenience samples (e.g. routinely collected sera from blood donors and pregnant women) would yield similar estimates of *Salmonella* seroincidence compared with community-based samples from healthy individuals since community-based samples are not widely available in most countries.

MATERIAL AND METHODS

Community cohort samples

In 2006, a random sample of 7770 Danish adults (with Danish citizenship and born in Denmark) aged 18–69 years living in one of the 11 municipalities of the former Copenhagen County were invited to participate in a general health examination. From June 2006 to July 2007, 3471 individuals agreed to participate. More

details about the procedures including non-response have been described previously [7]. All participants completed a questionnaire and had a blood sample taken. For 1780 randomly selected individuals, a serum sample was analysed for *Salmonella* antibodies.

Convenience samples

For the present study, two convenience samples were obtained from the Department of Clinical Immunology at Copenhagen University Hospital. The department tests sera from blood donors and pregnant women [8, 9] from about 35% of the Danish population, including the Capital region and Western Zealand. In October 2012, residual sera from 500 anonymous blood donors were collected and sent to Statens Serum Institut (SSI).

From August to December 2012, residual sera from pregnant women were forwarded to SSI aiming at the following age distribution: 80 sera per age group was the goal for all 5-year age groups from 20–24 up to 40–44 years while the aim was 50 sera for females aged <20 or >44 years.

ELISA method

All serum samples were tested for IgG, IgM and IgA *Salmonella* antibodies at the serological laboratory at SSI using an in-house *Salmonella* ELISA method, as described in Falkenhorst *et al.* [10].

Statistical analyses

The distribution of IgG, IgM and IgA, optical density (OD) values in blood donor sera, pregnancy sera and sera from the community cohort, were compared by the non-parametric Kruskal–Wallis test.

The IgG, IgM and IgA antibody responses were measured during a follow-up study of 303 patients with stool culture confirmed *S.* Typhimurium or *S.* Enteritidis infections [2, 10]. Based on these antibody measurements a mathematical model was developed to describe the longitudinal behaviour of the antibody response to a *Salmonella* infection. Using this model the measured antibody values from the cross-sectional samples were back-calculated to time since last infection and these values were converted to an incidence estimate. Details of this model can be found in Simonsen *et al.* [4].

Table 1. *Comparison of age distributions in the community cohort and in pregnant women*

Age group, yr	Community cohort No.	(%)	Women from community cohort No.	(%)	Men from community cohort No.	(%)	Pregnant women No.	(%)
<20	22	(1·2)	11	(1·1)	11	(1·4)	16	(2·5)
20–24	66	(3·7)	42	(4·3)	24	(3·0)	104	(16·3)
25–29	68	(3·8)	50	(5·1)	18	(2·3)	146	(22·9)
30–34	117	(6·6)	63	(6·4)	54	(6·8)	197	(30·9)
35–39	158	(8·9)	90	(9·2)	68	(8·5)	120	(18·8)
40–44	248	(13·9)	143	(14·6)	105	(13·1)	48	(7·5)
45–49	209	(11·7)	124	(12·6)	85	(10·6)	6	(0·9)
50–54	214	(12·0)	116	(11·8)	98	(12·3)		
55–59	237	(13·3)	127	(13·0)	110	(13·8)		
60–64	253	(14·2)	125	(12·7)	128	(16·0)		
65–71	188	(10·6)	90	(9·2)	98	(12·3)		
	1780		981		799		637	
Median age (yr)	50		48		51		31	
Interquartile range (yr)	40–59		39–58		41–60		26–35	

Seroincidences were compared pairwise by calculating means and percentiles of the posterior distributions of incidence rate ratios [5].

WinBUGS was used to estimate the longitudinal behaviour of antibodies after infection [11], while the statistical program SAS was used for the descriptive analyses and the estimation of incidence rates and ratios [12].

RESULTS

Descriptive statistics

Table 1 shows the age distributions of participants in the community cohort (all subjects as well as subjects stratified by gender), and of the serum collection of pregnant women. Sera from the community cohort originated from individuals with a median age of 50 years (range 18–71 years). The age distributions for males and females were similar. The proportion of females was 55·1%. The majority of females in the community cohort were aged >40 years while the majority of pregnant women were aged between 20 and 44 years (median age 31 years) (Table 1). The age and sex distribution was not available for the blood donors; however, our 500 sera came from a blood donor population with an age range from 17 to 67 years and a median age of 40 years. The sex distribution was 52·8% male and 47·2% female.

The measured IgG, IgM and IgA OD values were compared visually by box-and-whiskers plots (Fig. 1).

For all three antibody classes, the lowest OD values were observed in the samples from the community cohort whereas the highest values were observed in the samples from pregnant women. Testing for similar distributions using the Kruskal–Wallis test confirmed the observed IgG, IgM and IgA differences in pregnant women, blood donors and community cohort data. For each of the three antibodies IgG, IgM and IgA we found statistically different distributions between the three cohorts (P values <0·001).

Of the 637 pregnant women, 17 were aged <20 years and eight were aged ≥44 years. After omitting these from the analysis in order to obtain an age distribution similar to the community cohort, we compared the measured *Salmonella* IgG, IgM and IgA antibodies in 612 sera from pregnant women aged between 20 and 44 years with *Salmonella* antibodies in 371 sera from females aged between 20 and 44 years from the community cohort (Table 2). The antibody levels in pregnant women were still higher compared to women of the same age from the community cohort (Table 2).

Calculation of seroincidence

The estimated seroincidence in pregnant women was 319 infections/1000 population per year, corresponding to a mean annual risk of 27%. This mean seroincidence was 1·8 times higher than estimated from blood donors (182 seroconversions/1000, mean annual risk 17%). In the community cohort, the mean
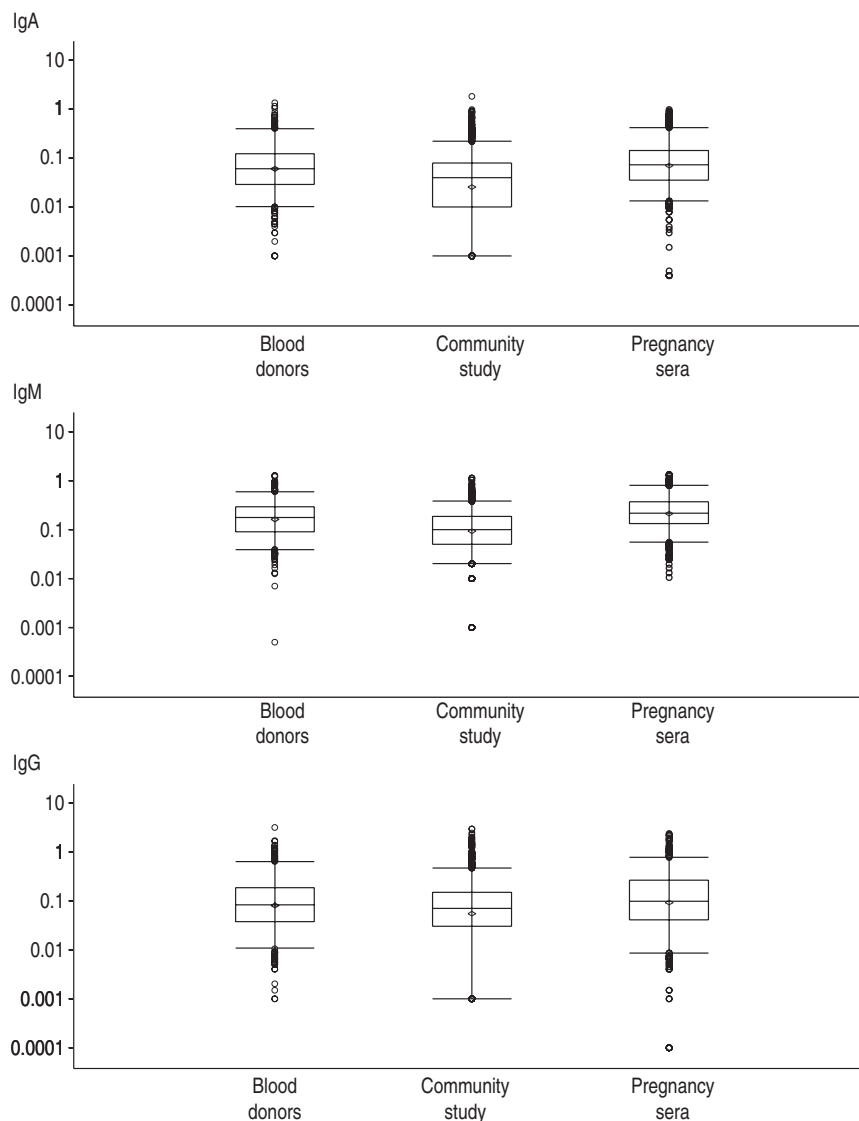
Fig. 1. The distribution of measured IgG, IgM and IgA optical density values in the three-sample community study, blood donors and pregnant women in a box-and-whiskers plot. The box represents the 25th–75th percentile, the line across the box is the median while the diamond is the mean, the lower and upper lines represent the 2·5th–92·5th percentiles and the circles represent the measurements below and above the lower and upper limits.

seroincidence was 77/1000, or a mean annual risk of 7%, which was 4·1 times lower than for pregnant women and 2·4 times lower than for blood donors (Table 3). The pairwise comparisons showed significantly different seroincidences between the community cohort and the pregnant women; however, the estimated seroincidences between the community cohort and the blood donors as well as between the pregnant women and the blood donors did not differ significantly at the 10% level (Table 4). For the community cohort, seroincidences were estimated for the following subgroups: males and females, participants

aged <44 years and ≥44 years, females aged <44 years and ≥44 years, and males aged <44 years and ≥44 years (Table 3). Pairwise comparisons between these subgroups showed no significant differences in seroincidences (Table 4); however, there is a tendency that younger individuals have a higher seroincidence compared to older individuals, particularly younger women (Table 3). We also estimated the seroincidence in women aged between 20 and 44 years participating in the community cohort and obtained an estimate of 128 seroconversions/1000; this was still significantly different compared to

Table 2. *Antibodies against non-typhoid* Salmonella *measured in mixed lipopolysaccharide ELISA as optical density (OD) values*

Populations	No. of samples	IgG (OD values)		IgM (OD values)		IgA (OD values)	
		Median	(25–75%)	Median	(25–75%)	Median	(25–75%)
Community cohort, 2006	1780	0·07	(0·03–0·15)	0·1	(0·05–0·19)	0·04	(0·01–0·08)
Blood donors, 2012	500	0·08	(0·04–0·18)	0·17	(0·09–0·29)	0·06	(0·03–0·12)
Pregnant women, 2012	637	0·1	(0·04–0·26)	0·22	(0·13–0·37)	0·07	(0·04–0·14)
Community cohort, 2006 (women only)	981	0·06	(0·03–0·14)	0·12	(0·07–0·20)	0·04	(0·01–0·08)
Community cohort, women 20–44 years	371	0·06	(0·02–0·16)	0·14	(0·08–0·26)	0·04	(0·01–0·08)
Pregnant women, 20–44 years	612	0·1	(0·04–0·26)	0·22	(0·14–0·37)	0·07	(0·04–0·15)

Median and interquartile range values are shown for the community study, blood donors, pregnant women and two subsamples including women aged between 20 and 44 years. The samples were collected in Denmark in 2006–2007 and 2012.

Table 3. *Estimated* Salmonella *seroincidence per year in Denmark, in a community cohort (years 2006–2007), in blood donors (year 2012) and in pregnant women (year 2012)*

Type of study	Participants	No. of samples	Estimated Salmonella seroincidence		Risk of at least one infection per year
			Mean	(90% CrI)	$1 - e^{(-(\text{mean seroinc})*1)}$
Community cohort	All participants	1780	0·077	(0·045–0·114)	~7%
	All women	981	0·082	(0·039–0·136)	~8%
	All men	799	0·068	(0·025–0·125)	~7%
	<44 years	643	0·097	(0·040–0·167)	~9%
	⩾44 years	1137	0·066	(0·030–0·111)	~6%
	Women <44 years	382	0·123	(0·045–0·225)	~12%
	Women ⩾44years	599	0·060	(0·014–0·124)	~6%
	Men <44 years	261	0·082	(0·010–0·194)	~8%
	Men ⩾44 years	538	0·064	(0·015–0·132)	~6%
	Women 20–44 years	371	0·128	(0·046–0·234)	~12%
Blood donors	All participants	500	0·182	(0·085–0·298)	~17%
Pregnant women	All participants	637	0·319	(0·210–0·441)	~27%
	20–44 years	612	0·307	(0·191–0·436)	~26%

CrI, Credibility interval.

pregnant women of the same age but the difference was not as marked as between all pregnant women and all women from the community (Table 4).

DISCUSSION

In many countries, programmes for *Salmonella* control have successfully applied serological methods to follow the occurrence of *Salmonella* in food animals [13–16]. There is currently no human counterpart to the systematic monitoring of food animals by serological methods. The end point for a reduction in

human illness is monitored by passive public health surveillance of culture-confirmed cases. It is well known that these data are subject to a number of limitations, including patterns of healthcare-seeking behaviour, practices at the healthcare level for obtaining faecal specimens for culture as well as laboratory and reporting practices [6, 17].

In European studies, we have demonstrated a close correlation between country-specific seroincidence and the occurrence of *Salmonella* in food animals as well as with a less biased assessment of the relative occurrence of *Salmonella* in humans (the 'Swedish

Table 4. *Pairwise comparisons of estimated* Salmonella *seroincidences using rate ratios*

Type of study	Comparisons	Rate ratios (90% CrI)
Community	Women *vs.* men	1·50 (0·46–3·55)
	Women <44 years *vs.* women ⩾ 44 years	2·83 (0·57–8·00)
	Men <44 years *vs.* men ⩾44 years	1·83 (0·14–5·93)
	Women <44 years *vs.* men <44 years	2·53 (0·40–8·26)
Community/pregnant	Community all *vs.* all pregnant women	0·26 (0·13–0·47)
Community/pregnant	Community women 20–44 years *vs.* pregnant women 20–44 years	0·44 (0·14–0·90)
Community/blood donors	Community all *vs.* all blood donors	0·50 (0·20–1·02)
Blood donors/pregnant	Blood donors *vs.* all pregnant women	0·63 (0·25–1·19)

CrI, Credibility interval.

travellers approach') [6, 18]. Most of these studies have taken advantage of collections of human samples that were carefully selected to achieve a high degree of representativeness of the general population. Such collections are difficult and costly to obtain, and it is important to be able to use more easily available serum collections to enable monitoring of the serological status in the human population as a feasible tool. In the present study we showed that *Salmonella* seroincidence in Denmark was more than four times higher when estimated from pregnant women compared to the community cohort. Although the difference between the seroincidence in pregnant women and individuals from the community cohort was reduced when the different age distributions in the two populations were taken into account we still found a significantly higher seroincidence in pregnant women compared to the community cohort. The samples were taken early in pregnancy and the outcome is likely to reflect a history of infections prior to conception, i.e. before pregnancy-related lifestyle changes may have modified the risk of infection. The individuals who participated in the community cohort had a higher educational level and a higher personal income than non-responders and non-responders were more often living alone than the responders [7]. It cannot be excluded that some participants who contributed to the community study did so because they have an interest in a healthy lifestyle, and therefore community-based samples may be more selected than desired by researchers.

The mean *Salmonella* seroincidence was 1·7 times higher when estimated from pregnant women than from blood donors; however, this comparison did not account for the fact that blood donor sera were drawn from a population with a median age of 40 years while pregnant women had a median age of 31 years. It is likely that a higher age distribution in blood donors compared to pregnant women has overestimated the difference in the mean seroincidence between pregnant women and blood donors and similarly between blood donors and the community cohort.

The finding of a higher seroincidence in pregnant women compared to the general population is in contrast with data from The Netherlands in 2006, where 530 leftover sera from the screening programme for pregnant women were analysed for *Salmonella* antibodies. The *Salmonella* seroincidence was 86/1000 (90% credibility interval 28–175) persons which tended to be lower than the population-based estimate of 149/1000 persons per year based on representative sera from 2006 to 2007 [19]. The reason for the observed difference in seroincidence in sera from pregnant women and population-based sera in The Netherlands and Denmark is not known. Both the Dutch and the Danish sera were analysed using the same in-house *Salmonella* ELISA from SSI and performed at SSI, and the same seroincidence back-calculation model was used for the seroincidence calculations. In addition, several technical reasons like handling and storage of sera have also been explored but no technical explanation can be identified.

The main limitation of the present study is the fact that the community-based samples were taken in 2006 and 2007 whereas our convenience samples were obtained 6 years later. As the number of notified human *Salmonella* cases in Denmark decreased from 1665 cases in 2006 to 1199 cases in 2012, a lower seroincidence in 2012 compared to 2006 might be expected. However, notified cases represent individuals so ill that they seek medical care. The estimated seroincidence represents all *Salmonella* infections including asymptomatic and mild symptomatic infections. Although fewer individuals become clinically ill from *Salmonella* this does not imply that a similar decrease in subclinical infections will occur.

The study is subject to other limitations. We know that the incidence of culture-confirmed *Salmonella* infections exhibits a marked seasonality, with more infections diagnosed in the late summer. The samples from the community cohort were collected during the entire year, whereas the convenience samples were collected in a short time period in the autumn of 2012. We have previously found that seasonality is a minor issue with regard to serology compared to the variation in stool cultures [20]. However, blood samples sent to SSI for diagnostic purposes during 2008–2013 were analysed for *Salmonella* antibodies and the first preliminary results indicated seasonality with the highest antibody levels in late summer. Therefore we cannot exclude that the season of convenience sampling introduced a bias towards a higher estimate compared to an all-season sampling scheme. Furthermore, children were not included in any of the studies.

Finally, we wish to emphasize that none of the estimated measures of seroincidence reflect the burden of clinical illnesses. This method does not distinguish between clinical infections and exposures to *Salmonella* that result in (re)activation of the immune system without apparent clinical illness.

In conclusion, we found that estimates of *Salmonella* seroincidence depend on the population from which the serum samples are drawn. Consequently, it is important to define a certain population under surveillance in order to apply serology in humans as a tool to monitor the progress of *Salmonella* control programmes in Europe. Anonymized residual blood sera where the age and sex distributions are known could potentially be used for such a monitoring programme. The impact of different age and sex distributions over time can be adjusted statistically. Although the results are likely to be valid when comparing within a population over time this needs to be further investigated before setting up a serological surveillance programme. Based on the mathematical model used in this study a seroincidence calculator tool was developed in the statistical program R by Peter Teunis from the Centre for Infectious Disease Control, National Institute of Public Health and the Environment (RIVM) in The Netherlands. From March 2015 this tool will be available from the ECDC homepage [21].

DECLARATION OF INTEREST

None.

REFERENCES

1. **EFSA, European Food Safety Authority, ECDC, European Centre for Disease Prevention and Control.** The European Union Summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2011 (www.efsa.europa.eu/efsajournal), 2013.
2. **Strid MA, et al.** Kinetics of the human antibody response against *Salmonella* enterica Serovars Enteritidis and Typhimurium determined by lipopolysaccharide enzyme-linked immunosorbent assay. *Clinical and Vaccine Immunology* 2007; **14**: 741–747.
3. **Kuhn KG, et al.** Detecting non-typhoid *Salmonella* in humans by ELISAs: a literature review. *Journal of Medical Microbiology* 2012; **61**: 1–7.
4. **Simonsen J, et al.** Estimation of incidences of infectious diseases based on antibody measurements. *Statistics in Medicine* 2009; **28**: 1882–1895.
5. **Simonsen J, et al.** Usefulness of seroconversion rates for comparing infection pressures between countries. *Epidemiology and Infection* 2011; **139**: 636–643.
6. **Falkenhorst G, et al.** Serological cross-sectional studies on *Salmonella* incidence in eight European countries: no correlation with incidence of reported cases. *BMC Public Health* 2012; **12**: 523.
7. **Thuesen BH, et al.** Cohort Profile: The Health 2006 cohort, Research Centre for Prevention and Health. *International Journal of Epidemiology* 2014; **43**: 568–575.
8. **Cowan S, et al.** Hepatitis B, HIV and syphilis screening of pregnant women, 2012. Statens Serum Institut,
9. **Blood donors in Denmark, annual 2012–2013** (http://www.bloddonor.dk/om-bloddonorerne/hvem-er-vi/aars-beretninger/Aarsberetning%202012.pdf) [in Danish]. Accessed 4 November 2013.
10. **Falkenhorst G, et al.** Serological follow-up after non-typhoid salmonella infection in humans using a mixed lipopolysaccharide ELISA. *International Journal of Medical Microbiology* 2013; **303**: 533–538.
11. **Lunn D, et al.** Winbugs – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 2000; **10**: 325–337.
12. **SAS.** SAS. Cary, NC, USA: SAS Institute.

13. **Wegener HC,** *et al.* *Salmonella* control programs in Denmark. *Emerging Infectious Diseases* 2003; **9**: 774–780.

14. **Merle R,** *et al.* Serological *Salmonella* monitoring in German pig herds: results of the years 2003–2008. *Preventive Veterinary Medicine* 2011; **99**: 229–233.

15. **Méroc E,** *et al.* Evaluation of the *Salmonella* surveillance program in Belgian pig farms. *Preventive Veterinary Medicine* 2012; **105**: 309–314.

16. **Stockmarr A, Bødker R, Nielsen LR.** Dynamic changes in antibody levels as an early warning of *Salmonella* Dublin in bovine dairy herds. *Journal of Dairy Science* 2013; **96**: 7558–7564.

17. **Hardnett FP,** *et al.* Epidemiologic issues in study design and data analysis related to FoodNet activities. *Clinical Infectious Diseases* 2004; **38**: S121–S126.

18. **Mølbak K,** *et al.* Seroincidence of human infections with non-typhoid *Salmonella* compared with data from public health surveillance and food animals in 13 European countries. *Clinical Infectious Diseases* 2014; **59**: 1599–1606.

19. **Med-Vet-Net Annual Report.** (http://www.mvnassociation.org/images/stories/documents/mvn_ar_year5.pdf), 2009.

20. **Simonsen J.** Sero-epidemiology as a tool to estimate the incidence of important and frequent infectious diseases. Faculty of Health, University of Copenhagen, 2011.

21. **Teunis PFM.** Seroincidence: estimating infection rates from serological data. ECDC homepage, 2015 (http://ecdc.europa.eu/en/data-tools/seroincidence-calculator-tool/Pages/default.aspx).