Epidemiol. Infect. (2014), **142**, 1070–1082. © Cambridge University Press 2013
doi:10.1017/S0950268813001829

Attribution of human *Salmonella* infections to animal and food sources in Italy (2002–2010): adaptations of the Dutch and modified Hald source attribution models

L. MUGHINI-GRAS[1,2]*, F. BARRUCCI[3], J. H. SMID[4], C. GRAZIANI[1], I. LUZZI[5], A. RICCI[3], L. BARCO[3], R. ROSMINI[2], A. H. HAVELAAR[4,6], W. VAN PELT[4] AND L. BUSANI[1]

[1] *Istituto Superiore di Sanità, Department of Veterinary Public Health and Food Safety, Rome, Italy*
[2] *Bologna University, Department of Veterinary Medical Sciences, Bologna, Italy*
[3] *Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro, Italy*
[4] *National Institute of Public Health and the Environment (RIVM), Centre for Infectious Disease Control, Bilthoven, The Netherlands*
[5] *Istituto Superiore di Sanità, Department of Infectious, Parasitic and Immunomediated Diseases, Rome, Italy*
[6] *Utrecht University, Institute for Risk Assessment Sciences, Utrecht, The Netherlands*

Received 10 March 2013; Final revision 22 May 2013; Accepted 9 July 2013; first published online 7 August 2013

SUMMARY

The Dutch and modified Hald source attribution models were adapted to Italian *Salmonella* data to attribute human infections caused by the top 30 serotypes between 2002 and 2010 to four putative sources (*Gallus gallus*, turkeys, pigs, ruminants), at the points of animal reservoir (farm), exposure (food), and both combined. Attribution estimates were thus compared between different models, time periods and sampling points. All models identified pigs as the main source of human salmonellosis in Italy, accounting for 43–60% of infections, followed by *G. gallus* (18–34%). Attributions to turkeys and ruminants were minor. An increasing temporal trend in attributions to pigs and a decreasing one in those to *G. gallus* was also observed. Although the outcomes of the two models applied at farm and food levels essentially agree, they can be refined once more information becomes available, providing valuable insights about potential targets along the production chain.

Key words: Italy, microbial subtyping, modelling, salmonellosis, source attribution.

INTRODUCTION

Salmonellosis is a major cause of human bacterial gastroenteritis and the second most reported foodborne zoonosis in the European Union (EU), after campylobacteriosis [1]. Humans can become infected with *Salmonella* from several sources and via different pathways, including direct contact with live animals, environmental and, to a lesser extent, anthroponotic transmission. However, the most common source is by far contaminated food, with 86–95% of cases estimated to be foodborne [2, 3]. In recent years, human cases of salmonellosis reported by Italian general practitioners have decreased markedly, falling from 47 to 7 cases/100 000 population in less than two decades [4]. This decrease has mainly concerned infections with *S.* Enteritidis, while infections with other serotypes have increased (e.g. *S.* Typhimurium monophasic variant 4,[5],12:i:- and *S.* Derby) or have remained fairly stable (e.g. *S.* Typhimurium) [5], suggesting that the relative importance of the different

* Author for correspondence: Dr L. Mughini-Gras, Department of Veterinary Public Health and Food Safety, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161, Rome, Italy.
(Email: lapo.mughinigras2@unibo.it)

sources of human salmonellosis has changed over time.

Attributing human *Salmonella* infections to specific sources is crucial to prioritize and implement targeted interventions in the food chain, as well as to measure the impact of such interventions [6]. The term 'source' is often used as a collective term to cover any point along the transmission pathway, such as the animal reservoirs or amplifying hosts (e.g. chicken, cattle, pig, etc.), the vehicles or exposures (e.g. food, water, direct contact with animals, etc.) and even specific food items (e.g. pork, milk, eggs, etc.). Several methods have been proposed for source attribution of foodborne diseases [7, 8]. In particular, the microbial subtyping approach, based on the comparison of the frequency distributions of pathogen subtypes isolated from humans with those isolated predominantly from putative animal, food and environmental sources, has received considerable attention since the development of the Hald model for *Salmonella* source attribution in Denmark [9]. The Hald model, a Bayesian adaptation of the earlier frequentist Dutch model [10], attributes stochastically human *Salmonella* infections to each putative source, to travelling abroad and to outbreaks, while accounting for differences in the different *Salmonella* subtypes and sources that cause human infection [9]. The Hald model has successfully been adapted to salmonellosis in several countries [6, 11–15]. Yet, to further improve its identifiability and to handle uncertainty in data of poorer quality, a modified Hald model has also been proposed [16].

While the Dutch model uses a straightforward approach, providing transparent insights into the functionality of the attribution process, the Hald model is a more complex model that fits parameters with no clear biological interpretation, and is therefore considered a sort of 'black box' model [11]. In absence of a universally agreed source attribution model to be used as the reference method, the comparative application of the Dutch and Hald models on the same data may be helpful in discerning the extent to which the obtained attributions are the result of the data used or that of the assumptions made by the models, providing valuable insights about how different methods may influence the attribution process.

So far, these two models have been applied to single points of the farm-to-food continuum, e.g. point of reservoir (food-producing animals), point of exposure (foods of animal origin), or both combined (undifferentiated). The comparative application of these two models to different points of attribution may further inform about the most promising targets on which risk management strategies should be focused. Indeed, the *Salmonella* serotype distribution within a source may well change from farm to food. Some serotypes may be predominant at every stage of the production chain while others may increase or decrease in importance due to, for example, their differences in survival. Moreover, other serotypes may enter the production chain, or pass from one chain to another, at later stages due to contamination from other sources. Thus, attributing human infections to sources on the basis of the *Salmonella* serotype distributions observed at farm and food levels separately, and at both these levels combined, is expected to provide insights about the relative importance of a set of putative sources to human infections along the farm-to-food continuum. This would allow identification of the most problematic points for *Salmonella* contamination in the different sources, as well as identifying which of these sources are the most problematic as a whole.

The main aim of this study was to adapt the Dutch and modified Hald source attribution models to Italian *Salmonella* data in order to estimate the proportions of domestic, sporadic human *Salmonella* infections attributable to four putative sources (*G. gallus*, turkeys, pigs, ruminants), which have been consistently monitored for a period of 9 years (2002–2010) both in animals and in foods of animal origin.

METHODS

Laboratory surveillance of *Salmonella* in humans

In Italy, testing for *Salmonella* infection is usually performed on patients with gastroenteritis seeking medical care or on people requiring periodic testing regardless of symptoms (e.g. food handlers, healthcare workers, etc.). Irrespective of symptomatology, *Salmonella* isolates from humans are reported to 'Enter-net Italia', a passive, laboratory-based surveillance system for human enteropathogens based on a network of more than 140 peripheral laboratories with about 65% coverage of the Italian territory, concentrated mainly on the northern part of the country. Enter-net Italia complements the Italian National Surveillance System for Infectious Diseases (SIMI) [17]. From the peripheral laboratories, Enter-net Italia collects demographical and microbiological information (at least the serotype) on *Salmonella*

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

isolates of ~50% of human cases of salmonellosis notified to SIMI [18]. Information on travel history or link to outbreaks concerns ~15% of serotyped isolates. At present, *Salmonella* isolates reported to Enter-net Italia are virtually indistinguishable between symptomatic and asymptomatic human infections. For the purposes of this study, a human *Salmonella* infection was considered to be: (1) travel-related if the person had travelled abroad during the 5 days before onset of symptoms; and (2) outbreak-related if the person had contacts with people with gastroenteritis and/or there have been other epidemiologically linked infections.

Veterinary surveillance of *Salmonella*

Findings of *Salmonella* in animals and foods of animal origin as part of diagnostic or monitoring activities are notifiable to Italian veterinary authorities. All major food-producing animals and foods of animal origin in Italy are tested for *Salmonella* according to EU-standardized official control programmes (Directive 2003/99/EC, Regulations EC 2160/2003 and 882/2004). For *G. gallus* and turkeys, sampling activities are implemented nationally in all farms, or in a representative percentage of them in broiler and fattening turkey farms, within the framework of official controls in order to meet EU targets for the reduction of *Salmonella* prevalence in poultry flocks as provided for by the specific EU legislation (Regulation EU No. 200/2010 for *G. gallus* breeders and broilers, Regulation EU No. 517/2011 for laying hens, and Regulation EU No. 1190/2012 for turkeys). Samples are processed for microbiological examination according to ISO method 6579:2002/Amd 1:2007 – Annex D, and the results are notified through a national web-based database. For pigs and ruminants, as well as for other animal species, no structured sampling schemes to detect the presence of *Salmonella* in farms are routinely applied at the national level; thus, pig and ruminant farm samples derive from locally implemented surveillance programmes, random controls or testing activities performed when there is suspicion of infection. In slaughterhouses, food processing plants and at retail, official samples for detection of *Salmonella* in foodstuffs are taken on a regular basis from all foods of animal origin, including those of the sources considered here, as verification of the application of Regulation EC 2073/2005. Positive results are then reported to 'Enter-vet', the Italian veterinary surveillance system for *Salmonella*. Enter-vet was established in 2002 and is based on a network of 10 peripheral laboratories covering the whole country through the regionally competent Institutes for Animal Health (Istituti Zooprofilattici Sperimentali), and is coordinated by the National Reference Laboratory for Salmonellosis. About 5000 *Salmonella* serotyped isolates from animals and foods of animal origin are reported to Enter-vet each year and classified by animal species and sampling point (farm or food). Enter-vet yearly data are made available online through periodical reports (http://www.izsvenezie.it/index.php?option=com_content&view=article&id=193&Itemid=335).

Salmonella data included in the models

The input dataset for the *Salmonella* attribution models included surveillance data over 9 years (from January 2002 to December 2010) collected by Enter-net and Enter-vet. Based on the most frequently isolated *Salmonella* serotypes from humans in common with at least one of the sources, the following 30 serotypes were included in the models: Typhimurium and its monophasic variant 4,[5],12:i:-, Enteritidis, Derby, Infantis, Muenchen, Hadar, London, Bredeney, Brandenburg, Rissen, Panama, Thompson, Virchow, Goldcoast, Give, Blockley, Newport, Heidelberg, Agona, Anatum, Saintpaul, Coeln, Montevideo, Kapemba, Mbandaka, Kedougou, Meleagridis, Senftenberg and Livingstone. The selected serotypes accounted for 20 890 human infections, corresponding to 87% of all human *Salmonella* infections reported in the study period. The remaining 13% of human infections caused by less frequent serotypes were excluded from the models and were not further considered in this study. A closer look at the data revealed that the excluded infections were often associated with travel and their serotypes were rarely, if ever, detected in the considered sources. Duplicate entries, i.e. different *Salmonella* isolates from a same person due to follow-up of patients with *Salmonella* infection after the first isolation, were discarded. This was done also for isolates belonging to the same serotypes derived from repeated sampling of the same farms, food processing plants or food batches at retail. Therefore, the models attributed only those human *Salmonella* infections that, during the entire study period and irrespective of clinical manifestations, were caused by the above-mentioned top 30 *Salmonella* serotypes found both in humans and in the

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

Table 1. *Parameters used to estimate the number of domestic and sporadic human* Salmonella *infections attributable to animal and food sources*

Notation	Description	Estimation
i	*Salmonella* serotype (30 serotypes)	Data
j	Animal or food source (4 sources)	Data
t	3-year period (2002–2004, 2005–2007, 2008–2010)	Data
o_{it}	Observed infections with serotype i in period t	Data
oyt_{it}	Observed infections with serotype i in period t reported to have travelled abroad in the incubation period	Data
ont_{it}	Observed infections with serotype i in period t reported to have not travelled abroad in the incubation period	Data
out_{it}	Observed infections with serotype i in period t with unknown travel history	Data
pt_{it}	Probability that a person infected with serotype i in period t with unknown travel history did travel	Beta($oyt_{it}+1$, $ont_{it}+1$)
et_{it}	Estimated number of additional infections with serotype i in period t that had travelled	Binomial(out_{it}, pt_{it})
dc_{it}	Estimated total number of domestic infections with serotype i in period t	$o_{it}-oyt_{it}-et_{it}$
oyb_{it}	Observed infections with serotype i in period t known to be outbreak-related	Data
oub_{it}	Observed infections with serotype i in period t with no information on relationships with outbreaks	Data
pb_{it}	Probability that a person infected with serotype i in period t is outbreak-related	Beta($oyb_{it}+1$, $out_{it}-oyb_{it}+1$)
eb_{it}	Estimated number of additional domestic infections with serotype i in period t that are outbreak-related	Binomial(dc_{it}, pb_{it})
e_{it}	Estimated total number of domestic and sporadic infections with serotype i in period t	$dc_{it}-oyb_{it}-eb_{it}$

considered animal and food sources. The specific conditions for travel- and outbreak-related human *Salmonella* infections listed above were used to identify those human infections that were known to have been acquired abroad or be part of outbreaks within the Enter-net dataset prior to attribution analysis. The number of human infections with unknown history of travel or unknown involvement in outbreaks that had actually travelled abroad or had actually been involved in outbreaks was then estimated as reported in Table 1 so that the models only attributed to sources domestic and sporadic human infections. Domestic and sporadic human infections were therefore defined as infections acquired in Italy and not implicated in outbreaks.

Frequencies of human infections were merged with animal and food isolates by serotype, sampling point and year. Based on data availability, the following sources were considered: *G. gallus*, turkeys, pigs, and ruminants (cattle, sheep and goats combined). These sources were consistently sampled at the farm level (live animal faecal samples and boot sock swabs) and at retail (foods of animal origin) during the entire study period. Differentiation of *G. gallus* between broilers and layers/eggs was not possible because the data were available at the species level only.

To avoid sparse data that may lead to a low precision of serotype prevalence estimates [16], the merged dataset was arranged in three 3-year periods (2002–2004, 2005–2007, 2008–2010). This arrangement was made for temporal ordination purposes only, with no direct correspondence with the sampling schemes or reporting procedures, which were uniform and consistent over the entire study period in both humans and sources. The resolution of phage-typing data was very low and did not allow for the use of this information in the analysis. Serotype frequencies in humans, animal and food sources are reported in Table 2.

Overview of the models

A modified version of the Dutch model and a Hald model accommodating for temporal dimension [11], with some further adjustments as proposed by Mullner *et al*. [16], were developed to estimate the proportions of domestic, sporadic human *Salmonella* infections in Italy attributable to the four putative sources at farm (animal reservoir) level, at food (exposure) level, and at both levels combined.

Where the 95% credible intervals (CrIs) of the attribution estimates did not overlap each other,

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

Table 2. *Frequencies of* Salmonella *serotypes isolated from humans and from animal and food sources, at farm and food levels, in (I) 2002–2004, (II) 2005–2007, and (III) 2008–2010, Italy*

				Gallus gallus						Pigs						Turkeys						Ruminants					
	Humans			Farm			Food			Farm			Food			Farm			Food			Farm			Food		
Serotype	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
Typhimurium	3140	2667	2919	129	188	161	45	274	35	456	535	371	796	502	398	108	195	10	73	60	52	199	160	175	90	112	77
4,[5],12:i:-	136	300	1324	9	9	74	6	88	22	138	263	817	106	175	609	4	11	16	4	5	24	5	28	100	4	12	35
Enteritidis	2181	1453	1212	159	244	377	167	100	82	1	16	3	10	159	8	1	1	6	1	5	28	3	5	6	11	6	20
Derby	239	253	344	5	6	8	16	159	17	159	359	164	577	310	331	26	14	4	20	8	10	6	16	22	26	12	13
Infantis	245	232	185	40	31	60	47	30	23	6	23	41	99	63	32	1	1	0	0	0	12	1	2	4	2	1	9
Muenchen	144	67	145	0	24	193	2	5	44	0	2	14	22	19	3	0	1	0	1	1	10	0	0	3	0	4	2
Hadar	141	60	127	187	148	224	215	50	93	3	8	2	12	48	2	59	16	7	46	19	41	6	0	1	9	3	2
Rissen	54	52	124	0	6	13	10	32	5	0	85	46	77	76	93	0	1	0	0	0	0	0	2	1	4	5	30
London	103	61	103	8	0	5	32	39	4	9	36	61	139	66	67	1	0	0	1	0	0	3	2	5	8	0	9
Bredeney	108	60	96	0	0	0	25	53	90	0	0	0	124	116	24	0	0	0	10	41	31	0	0	0	12	7	5
Newport	36	41	96	0	0	15	0	0	12	0	0	1	0	0	5	0	0	74	0	0	95	0	0	9	0	0	2
Goldcoast	74	30	89	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	1	0	0	0	0	0	1	0	0
Brandenburg	97	58	84	132	101	161	1	1	0	53	44	26	67	27	0	3	103	22	0	0	0	6	22	4	1	1	0
Give	44	71	74	3	0	1	0	29	0	2	19	11	17	0	14	0	0	0	1	0	0	15	1	7	0	3	0
Panama	110	40	72	0	0	0	0	0	0	0	0	0	46	0	0	0	0	0	0	0	0	0	0	0	4	0	0
Thompson	79	68	62	63	50	186	29	24	19	1	11	2	1	65	1	1	1	0	0	1	0	5	1	8	3	0	2
Coeln	11	15	58	0	2	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0
Agona	61	30	50	13	13	61	27	14	0	0	4	3	19	10	0	36	26	1	22	34	0	1	0	1	5	8	0
Saintpaul	60	22	48	5	2	0	50	1	19	0	1	0	6	1	6	13	22	0	39	6	58	1	1	0	10	1	4
Virchow	89	60	46	256	135	0	68	1	0	4	0	0	4	51	0	1	0	0	1	0	0	4	0	0	1	2	0
Anatum	54	44	41	10	3	14	8	73	0	67	54	21	123	40	0	33	4	0	41	7	0	3	4	3	14	4	0
Livingstone	71	47	39	0	0	0	129	93	21	0	0	0	73	192	5	0	0	0	0	2	9	0	0	0	2	3	1
Kapemba	9	13	38	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	4
Blockley	118	34	22	62	34	13	56	26	0	9	1	1	8	21	0	148	40	5	99	27	0	8	3	1	9	1	0
Montevideo	24	27	21	0	42	110	9	0	45	0	1	0	1	0	7	0	0	5	1	0	0	0	0	2	1	0	1
Heidelberg	118	27	18	109	42	96	64	16	0	8	8	0	5	37	0	92	143	1	45	115	0	1	3	0	4	4	0
Mbandaka	10	3	12	0	37	129	0	15	0	0	1	0	0	58	0	0	0	0	0	1	0	0	2	0	0	1	1
Kedougou	9	2	9	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0
Meleagridis	9	4	5	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Senftenberg	5	5	2	23	19	22	12	0	3	0	1	0	3	0	0	2	1	1	0	0	3	0	0	2	1	0	2
Total	7579	5846	7465	1213	1136	1935	1018	1123	535	916	1472	1584	2372	2036	1626	529	580	152	406	332	390	267	252	354	222	190	219

these were considered to be significantly different from one another at the 5% level of significance.

Modified Dutch model

The original Dutch model compares the number of human *Salmonella* infections caused by a particular serotype with the relative occurrence of that serotype in each source [10]. The expected number of human infections (λ_{ijt}) caused by serotype i from source j in period t is given by:

$$\lambda_{ijt} = \frac{r_{ijt}}{\sum_j r_{ijt}} \times e_{it},$$

where r_{ijt} is the relative occurrence of serotype i from source j in period t, and e_{it} is the estimated number of sporadic and domestic human infections of serotype i in period t (see Table 1 for notations and estimation of e_{it}). A sum over serotypes gives the total number of infections expected from source j in period t, denoted by: $\lambda_{jt} = \sum_i \lambda_{ijt}$.

In this study, the Dutch model was modified to incorporate prevalence uncertainty and food consumption weights. Prevalence was modelled using the novel approach proposed by Mullner *et al.* [16] based on the assumption that $p_{ijt} = \pi_j \times r_{ijt}$, where p_{ijt} is the prevalence of serotype i from source j in period t, π_j is the overall prevalence of all *Salmonella* serotypes in source j, and r_{ijt} is the relative occurrence of serotype i from source j in period t. Uncertainty was introduced in the estimates of the prevalence by assuming the following probability distributions:

$$\left(r_{1jt}, r_{2jt}, ..., 1 - \sum_{i=1}^{I-1} r_{ijt}\right) \sim \text{Dirichlet}(X_{1jt}, X_{2jt}, ..., X_{Ijt}),$$

where X_{ijt} (with $i = 1, 2, \ldots, I$) are the source isolates of serotypes i from source j at time t, and $\pi_j \sim \text{Beta}(\alpha_j + 1, \beta_j + 1)$, where α_j are the *Salmonella*-positive sampling units from source j and $\beta_j = N - \alpha_j$, with N being the total number of sampling units from source j that have been tested for *Salmonella* spp. The number of tested sampling units and respective positivity percentages in different animal reservoirs in Italy were provided by Pires *et al.* [14] by collating available information from the EU *Salmonella* prevalence baseline survey and from the EU Summary Reports on Trends and Sources of Zoonoses, Zoonotic Agents and Food-Borne Outbreaks, as published annually by the European Food Safety Authority from 2006 to 2009. These data were provided at animal/sample levels for broilers, bovines and pigs, and at flock/herd levels for layers and turkeys. Priors for parameter π_j were the same for both models applied at farm and food levels and for all time periods (pigs: $\alpha_j = 116$, $N = 709$; *G. gallus:* $\alpha_j = 237$, $N = 1214$; turkeys: $\alpha_j = 277$, $N = 1370$; ruminants: $\alpha_j = 17$, $N = 1797$).

Average *per capita* daily food consumption (g/person per day) for source j in period t in Italy, denoted as m_{jt}, was obtained from the Eurostat database (http://epp.eurostat.ec.europa.eu/portal/page/portal/food/data/database) for ruminant and pig meats. As the Eurostat database provides data on poultry consumption as a whole with no differentiation between *G. gallus* (meat/eggs) and turkey, we used the data from the Italian National Association of Poultry Producers (UNA; http://www.unionenazionaleavicoltura.it/prodcons.aspx). Uncertainty was introduced in the estimates of m_{jt} by assuming that $\log(m_{jt}) \sim \text{Normal}(\mu_{jt}, \sigma_{jt})$, where μ_{jt} and σ_{jt} are respectively the mean and standard deviation of the *per capita* daily food consumption for source j in period t. Both μ_{jt} and σ_{jt} were computed over the three annual values falling within the periods 2002–2004 (pigs: $\mu_{jt} = 106{\cdot}4$, $\sigma_{jt} = 1{\cdot}3$; *G. gallus:* $\mu_{jt} = 73{\cdot}9$, $\sigma_{jt} = 1{\cdot}4$; turkeys: $\mu_{jt} = 12{\cdot}7$, $\sigma_{jt} = 0{\cdot}4$; ruminants: $\mu_{jt} = 71{\cdot}2$, $\sigma_{jt} = 0{\cdot}9$), 2005–2007 (pigs: $\mu_{jt} = 106{\cdot}5$, $\sigma_{jt} = 3{\cdot}3$; *G. gallus:* $\mu_{jt} = 69{\cdot}8$, $\sigma_{jt} = 1{\cdot}9$; turkeys: $\mu_{jt} = 10{\cdot}8$, $\sigma_{jt} = 0{\cdot}3$; ruminants: $\mu_{jt} = 71{\cdot}7$, $\sigma_{jt} = 0{\cdot}9$), and 2008–2010 (pigs: $\mu_{jt} = 103{\cdot}2$, $\sigma_{jt} = 1{\cdot}6$; *G. gallus:* $\mu_{jt} = 72{\cdot}9$, $\sigma_{jt} = 0{\cdot}4$; turkeys: $\mu_{jt} = 11{\cdot}2$, $\sigma_{jt} = 0{\cdot}4$; ruminants: $\mu_{jt} = 65{\cdot}8$, $\sigma_{jt} = 1{\cdot}0$).

Using the above notations and those in Table 1, the modified Dutch model we used is denoted by:

$$\lambda_{ijt} = \frac{p_{ijt} \times m_{jt}}{\sum_j p_{ijt} \times m_{jt}} \times e_{it},$$

The model was implemented in @Risk (Palisade Corp., USA) by setting 100 000 iterations with the Latin hypercube sampling technique and a seed of 1. The final formula of the modified Dutch model resulted as being substantially similar to that of the 'Simple Attribution Model' proposed by David *et al.* [15], but the estimation of each of the parameters was slightly different.

Modified Hald model

The Hald model compares the number of human infections caused by different serotypes with their prevalence in the different sources, accounting for the amount of food consumed and incorporating serotype- and source-dependent factors [9]. By using a Bayesian approach, this model can explicitly incorporate prior information and quantify the uncertainty

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

around each of the parameters. We applied the modified version of the Hald model as described elsewhere [16]. Using the above notations and those reported in Table 1, we assumed that:

$$o_{it} \sim \text{Poisson}\left(\sum_j \lambda_{ijt}\right),$$

and that

$$\lambda_{ijt} = m_{jt} \times p_{ijt} \times q_i \times a_j,$$

where o_{it} is assumed to be Poisson distributed; p_{ijt} was modelled using the above-mentioned novel approach of Mullner *et al*. [16]; q_i is the serotype-dependent factor, which putatively accounts for differences in survivability, virulence and pathogenicity of serotypes *i*; and a_j is the source-dependent factor, which putatively accounts for the ability of the sources *j* to act as vehicles for *Salmonella* (e.g. differences in pathogen load, source characteristics influencing pathogen growth, preparation/handling procedures, differences in sensitivity of surveillance programmes and randomness of sampling schemes). In accordance with Mullner *et al*. [16], both q_i and a_j were assumed to be constant over time and q_i was modelled hierarchically as $\log(q_i) \sim \text{Normal}(0, \tau)$, where τ is given by a fairly diffuse Gamma(0·01, 0·01) distribution. Parameter a_j was defined as uninformative uniform (0, 100) distribution. Parameter q_i for *S*. Typhimurium monophasic variant 4,[5],12:i:- was set to be equal to that of *S*. Typhimurium. However, exploratory analyses revealed that setting different q_i parameters for *S*. Typhimurium and its monophasic variant 4,[5],12:i:- had no influence on model results.

Posterior distribution was obtained by a Markov Chain Monte Carlo simulation implemented in WinBUGS 1.4 (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml). Five independent Markov chains were run for 30 000 iterations after a burn-in period of 10 000 iterations, which was able to provide convergence as monitored by the method developed by Gelman & Rubin [19].

RESULTS

Modified Dutch model

Mean percentages and respective 95% CrIs of human *Salmonella* infections attributed to each of the sources, to travelling abroad and to outbreaks by the modified Dutch model are reported by time period in Figure 1. Overall (2002–2010), pigs were the source causing the highest percentage of human *Salmonella* infections attributed to animals (43%, 95% CrI 42–44), food (45%, 95% CrI 44–46) and both combined (44%, 95% CrI 43–45), followed by *G. gallus* (farm: 34%, 95% CrI 32–35; food: 32%, 95% CrI 31–33; farm + food: 33%, 95% CrI 32–34); turkey (4%, 95% CrI 4–5 at all levels) and ruminants (2%, 95% CrI 2–3 at all levels). Infections estimated to be travel- and outbreak-related amounted to 16% (95% CrI 15–17) and 1% (95% CrI 1–1), respectively.

A significant decrease in the percentage of infections attributed to *G. gallus* was observed from 2002–2004 to 2008–2010 (−6%, −4% and −4% on average, for each 3-year period at the levels of farm, food and both combined, respectively), whereas the percentage of infections attributed to pigs increased significantly (+4%, +2% and +3% on average, for each 3-year period at the levels of farm, food and both combined, respectively). Percentages of infections attributed to other sources, to travelling abroad and to outbreaks did not vary significantly over time (Fig. 1).

Modified Hald model

Percentages of human *Salmonella* infections attributed to each of the sources, to travelling abroad and to outbreaks by the modified Hald model are reported by time period in Figure 1. Pigs were again the source that accounted for the highest percentage of infections attributed to animals (60%, 95% CrI 48–72%), food (47%, 95% CrI 41–52) and both combined (47%, 95% CrI 42–52), followed by *G. gallus* (farm: 18%, 95% CrI 4–31; food: 33%, 95% CrI 28–38; farm + food: 32%, 95% CrI 27–37). Turkeys were the third most important source at the farm level (3%, 95% CrI 0–7) and at both farm and food levels combined (2%, 95% CrI 0–5), but was fourth at the food level (1%, 95% CrI 0–4), behind ruminants (farm: 2%, 95% CrI 0–5; food: 1%, 95% CrI 0–3; farm+food: 1%, 95% CrI 0–3). Infections estimated to be travel- and outbreak-related amounted to 16% (95% CrI 15–17) and 1% (95% CrI 1–1), respectively.

From 2002–2004 to 2008–2010, percentages of infections attributed to *G. gallus* decreased by −4% (animals), −5% (food) and −5% (both animals and food combined) on average, for each 3-year period, whereas those attributed to pigs increased by +2% (animals), +2% (food) and +4% (both animals and food combined). However, none of these trends was significant as the CrIs of attribution estimates were largely overlapping. Percentage of cases attributed

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

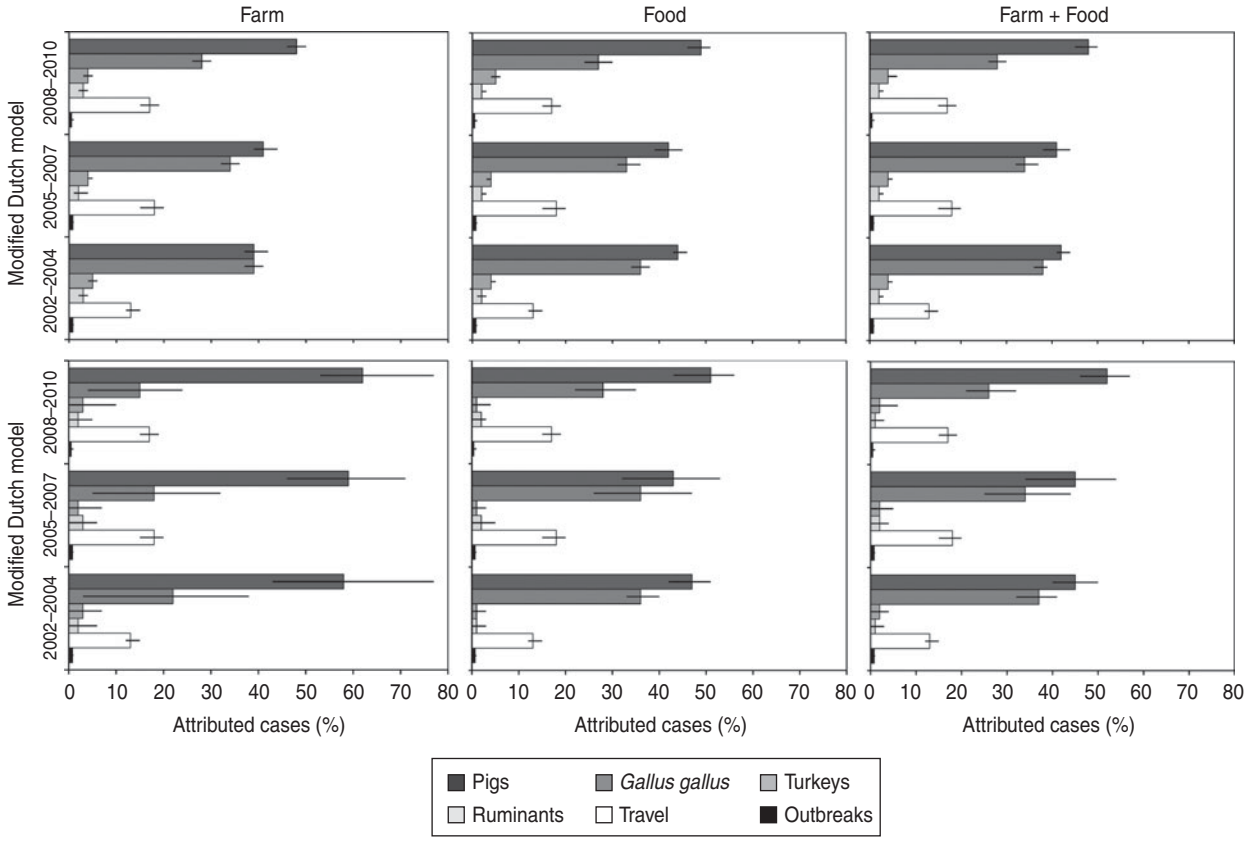

Fig. 1. Percentages of human *Salmonella* infections attributed to each putative animal (farm) and/or food source, to travelling abroad and to outbreaks, estimated by the modified Dutch and Hald models.

to other sources, to travelling abroad and to outbreaks did not vary significantly over time (Fig. 1).

DISCUSSION

In this study, two models were developed to attribute domestic and sporadic human *Salmonella* infections caused by the 30 most frequently reported serotypes in Italy between 2002 and 2010 to four putative sources at the points of reservoir (food-producing animals), exposure (foods of animal origin), and both combined. This allowed us to compare the obtained attribution estimates between different models, sampling points and time periods.

Pigs stood out as the largest contributor to human salmonellosis in Italy. This finding was consistent over different models, time periods and sampling points, and was also in line with previous estimates based on the (original) Hald model applied to a rather different input dataset in which 73% of human *Salmonella* infections that occurred in Italy between 2007 and 2009 had indeed been attributed to pigs [14].

Besides Italy, another seven (out of 24) European countries considered by Pires *et al.* [14] have identified pigs as the most important source of human salmonellosis. These countries were Belgium, Cyprus, Finland, France, Ireland, Poland and Sweden, with very similar proportions of infections attributed to poultry and to pigs in The Netherlands [14]. Moreover, in New Zealand pigs have been identified as the major source of human salmonellosis, accounting for 60% of infections, followed by poultry [16]. It is therefore increasingly evident that pigs play a paramount role as source of human salmonellosis, at least in the EU, and that (mis)handling and consumption of contaminated pork is the most likely foodborne pathway involved.

We observed an increasing temporal trend in the percentages of infections attributed to pigs and a concurrent reduction of those attributed to *G. gallus*. The decreasing importance of *G. gallus* is mainly driven by the marked decrease in the number of human *S.* Enteritidis infections (Table 2, Fig. 2), for which *G. gallus*, and particularly layers, are the major reservoir [14, 20]. Such decrease has been observed in

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

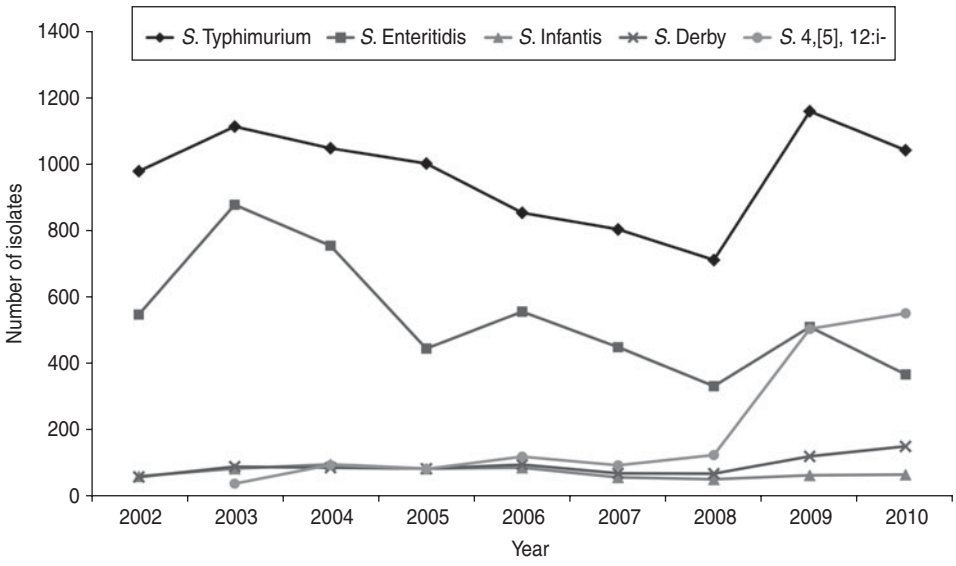

Fig. 2. Inter-annual trends of the top five *Salmonella* serotypes isolated from humans during 2002 to 2010.

most European countries, including Italy, since the late 1990s as a result of the implementation of new on-farm control measures in poultry (e.g. introduction of live vaccines), improved hygiene, and education of consumers and food workers [14, 21–24], especially after the implementation of national control programmes for *Salmonella* in poultry according to EU Regulation (EC) No. 2160/2003. Conversely, the increasing importance of pigs is mainly driven by the predominance of human infections caused by *S.* Typhimurium and its monophasic variant 4,[5],12:i:-, as well as by the increase of those caused by *S.* Derby (Table 2, Fig. 2). Indeed, pigs are the most likely reservoir for *S.* Typhimurium, its monophasic variant 4,[5],12:i:- and *S.* Derby [14, 20], and since 2000 in Italy, human *S.* Enteritidis infections fell consistently below those caused by *S.* Typhimurium, which has therefore become the most frequently isolated serotype from humans [5].

In all periods and sampling points, the two models have identified turkeys and ruminants as minor sources, accounting for 1–5% of human *Salmonella* infections. This is in line with previous estimates indicating that ~3% of all human *Salmonella* infections in the EU are attributable to turkeys relative to broilers, layers and pigs [6]. Ruminants have seldom been included as a putative source in attribution studies conducted in the EU, mainly because of data availability issues [14]. Although ground beef seems to be an important source of human salmonellosis in the USA [13], there is also some evidence indicating that ruminants do not play such a significant role [11, 12, 16].

Both our model adaptations retained much of the original methodology. In its application the Dutch model relies on a relatively straightforward frequentist approach, and it is possibly the easiest to run and interpret. However, as stated in the Introduction, this model has the major disadvantage of not explicitly accounting for differences in serotypes and sources to cause human infection, and it may therefore be argued that it is too simplistic. In contrast, the Hald model does explicitly include both serotype- and source-dependent factors and incorporates parameter uncertainty using Bayesian inference. However, this model is computationally demanding, tends to be over-parameterized and it is prone to non-convergence [16].

Modelling the prevalence using the novel methodology of Mullner *et al.* [16] allowed us to take into account the overall probability of finding *Salmonella* in a given source (parameter π_j) in addition to the relative frequency of the different serotypes within each source (parameter r_{ijt}, reflecting our best guess of the within-source serotype probability distribution). This is a necessary step towards compensating for the absence of intensive surveillance data for all relevant sources as required by the original Hald model [9]. Moreover, uncertainty around such estimates could not be ignored without overestimating the level of precision [16]. Therefore, by incorporating this additional stratum of information and uncertainty, the models

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

can now make use of the best possible estimate of the prevalence. Nevertheless, concerns remain about the adequacy of the priors used for modelling the parameter π_j, as these originated from both individual- and flock-level sampling schemes [14] and did not change either over attribution points or over time periods. This implied that the overall probability of finding *Salmonella* in a given source, as expressed by pooling the available data at different resolutions, was assumed to be a property of the sources themselves and to be relatively stable over time and along the farm-to-food continuum. Changes in prevalence were therefore primarily due to changes in within-source serotype distribution.

As reliable food consumption data were available and environmental, anthroponotic or unknown sources were not included in the models, food consumption weights were incorporated to take into account human exposure to the different sources. The importance for human *Salmonella* infection of foodborne exposure is unquestionable [2, 3]; thus, by incorporating food consumption data the models are better informed and can more closely reflect the chance of a given source to act as a vehicle for *Salmonella*. This incorporation is particularly relevant in the modified Dutch model because this model does not assume that within each subtype the impact of the different sources is equal and proportional to r_{ijt} only, as sources taking higher π_j and m_{jt} values can now result in more infections attributed to that source. The amount of food consumed, however, does not in itself inform the models about how this food was stored, prepared or consumed. Some sources may in fact pose a greater risk than others due to, for example, their greater likelihood to be consumed raw or undercooked (e.g. eggs and beef). Thus, food consumption data cannot in themselves account entirely for the chance of a source to act as a vehicle for *Salmonella*. It follows, therefore, that because the modified Dutch model does not include the Hald model's parameter a_j accounting for those source-dependent characteristics not accounted for by the amount of food consumed, further (and often subjective) considerations are necessary for correctly interpreting the impact of the different sources to human infections [9].

Attribution estimates of the modified Dutch model seem to be more precise and consistent between farm and food levels compared to those of our modified Hald model, which seem to be more sensitive to changes in within-source serotype frequency distribution between farm and food levels. Discrepancies in the estimates between the two models may be explained by the different computational methods they use, as also pointed out elsewhere [15, 16, 25]. Specifically, estimates from our modified Hald models were associated with large uncertainty. While this seems to be a characteristic of the modified Hald model [16, 25], it is somewhat offset when combining farm and food data, a reflection of decreased heterogeneity of serotypes between sources when models were developed for farm and food individually.

A heterogeneous distribution of the frequently occurring serotypes in the sources is a prerequisite for the Hald model to find the solution with the highest probability of occurrence. Violating such heterogeneous distribution would result in a very diffuse posterior distribution, as the frequency of the so-called 'indicator serotypes' on which source attribution relies would be of little information for the model [9]. In our modified Hald models, although convergence was adequately achieved, we observed signs of this, as the distributions of the infections attributable to *G. gallus* and pigs were rather wide, especially at the farm level. In particular, attribution estimates drifted from *G. gallus* to pigs at the farm level, and away from pigs and turkeys at the food level, thereby allowing the contribution of *G. gallus* to human salmonellosis to increase considerably from farm to food. This may be due to the fact that serotypes predominating in *G. gallus* and pigs (at least in animals) were also frequently found in other sources (Table 2), but is also suggestive of an important role of hygiene practices in modifying the within-source serotype distribution along the food production chain in such a way that the contribution of *G. gallus* to human infections as estimated using the serotype frequencies at the food level is higher than that at the farm level. As the priors for the overall prevalence of *Salmonella* in the sources were kept invariant in the models, differences in attributions from farm to food were essentially driven by changes in the within-source serotype distributions, and these changes may be the results of several factors, including serotype-specific survival abilities, selective pressures on the contaminating *Salmonella* populations induced by the changing environment and food processing procedures, contamination from other sources, and sampling schemes that may be more or less representative of certain serotypes. While attribution at the point of production would identify the most problematic animal reservoirs due to on-farm *Salmonella* contamination prior to or

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

during harvesting (i.e. contamination as observed in the animals themselves that can be transferred to carcases upon slaughtering), attribution at the point of exposure would identify the most problematic foods as they are prepared and eaten. Thus, in human risk terms, the most problematic point for *Salmonella* contamination in the *G. gallus* production chain seems to be that of food relative to that of farm, and the inverse holds for pigs.

High sensitivity of the Hald model to changes in prior information, particularly for the serotype-dependent parameter q_i, has been claimed [15]. We chose to model q_i hierarchically as a random effect with its variation controlled by the hyper-parameter τ, as in the modified Hald model [16]. This, together with the use of data split into multiple periods while estimating pooled q_i and a_j parameters over all periods, was expected to improve the identifiability and robustness of the model, as reported elsewhere [11, 16]. Inherent to the way by which these parameters were estimated is the assumption that the ability of the different serotypes and sources to cause infection in humans are properties of the serotypes and of the sources themselves, and these do not change over time. Temporal differences were therefore expected to be explained entirely by the serotype frequency distributions, food consumption patterns and sampling uncertainty. Some of the same considerations as discussed for the source-dependent factor also apply when a serotype-dependent factor is not included in the (modified) Dutch model. Indeed, Hald model's parameter q_i summarizes a complex system that is still not fully understood [11]; besides differences in survivability, virulence and pathogenicity, other factors such as undetected outbreaks can affect q_i values [12]. It would therefore be appropriate for future source attribution models to be better informed with exogenous data on the infectious capacity of the different serotypes.

Attributions made here have some limitations related to data availability and are in need of further investigation. These included the lack of distinction between broilers and layers/eggs within *G. gallus* and the lack of more discriminatory typing data than serotypes only. Furthermore, concerns remain about the heterogeneous distribution of human *Salmonella* infections across the country, as southern regions are usually more prone to underreporting than northern regions [4, 5]. It should also be borne in mind that, in contrast to *Salmonella* surveillance in foodstuffs and in poultry farms, there is no nationally structured sampling scheme in Italy for the detection of *Salmonella* in pigs and ruminants at the farm level. However, this appears to have had no apparent effect on our attributions, at least in the ranking of sources, as both our models applied at farm and food levels were consistent in identifying pigs as the most important source of human salmonellosis, and ruminants as the least important one. This is probably a reflection of the fact that the occurrence of clinical salmonellosis in pigs and ruminants, at least in Italy, is very rare, so the impact of diagnostic samples on the within-source serotype distributions of the Enter-vet dataset was negligible. Diagnostic samples are indeed mainly taken from pets and other animals primarily kept for leisure activities than for food production (e.g. equines) that were not considered in this study.

As a final point, we only considered four major (foodborne) sources of human salmonellosis of animal origin. A recently identified trend in human salmonellosis has been an increased association with unusual vehicles that are not routinely monitored for *Salmonella* in Italy, such as fresh produce [26] and low-moisture foodstuffs (e.g. peanut butter, infant formula, chocolate, cereal products, dried milk) [27]. Although these foods may be responsible for some of the infections in the Enter-net dataset, their omission from the models is common in *Salmonella* source attribution studies and is generally considered to be acceptable as long as most primary contaminations are assumed to originate from an animal reservoir (e.g. vegetables contaminated with animal manure used as fertilizer) or are identified as part of outbreaks [9].

CONCLUSIONS

With some differences in consistency and precision of attribution estimates over time periods and sampling points, both our adaptations of the modified Dutch and Hald source attribution models identified pigs as the main source of human salmonellosis in Italy, followed by *G. gallus*, whereas the contributions of turkeys and ruminants were estimated to be only minor. This ranking provided us with valuable insights about the relative contribution of these sources to the burden of human salmonellosis in Italy. The increasing importance of pigs and the decreasing importance of *G. gallus* as sources of human salmonellosis suggest that the applied control measures have been successful in poultry but there is an urgent need to focus

https://doi.org/10.1017/S0950268813001829 Published online by Cambridge University Press

attention on pigs. Despite data limitations and uncertainty in the results, our attribution estimates can be considered valid as a first indication of which sources are becoming increasingly important in Italy. These results are expected to be useful for the delineation of future risk management strategies in Italy. Although both our models applied at farm and food levels reached similar conclusions, more detailed data collected at varying levels of the transmission chain may further inform policy makers about the most critical points on which control efforts should be targeted.

DECLARATION OF INTEREST

None.

REFERENCES

1. **European Food Safety Authority, 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 2010. *EFSA Journal* 2012; **10**: 2597.
2. **Scallan E, *et al.*** Foodborne illness acquired in the United States – major pathogens. *Emerging Infectious Diseases* 2011; **17**: 7–15.
3. **Majowicz S, *et al.*** The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clinical Infectious Diseases* 2010; **50**: 882–889.
4. **Mughini-Gras L, *et al.*** Surveillance of acute infectious gastroenteritis (1992–2009) and food-borne disease outbreaks (1996–2009) in Italy, with a focus on the Piedmont and Lombardy regions. *Eurosurveillance* 2012**; 17**: pii=20098.
5. **Dionisi AM, *et al.*** Enter-net: surveillance of infections transmitted by food and water. 2007-2009 activity report [in Italian]. *Notiziario dell'Istituto Superiore di Sanità* 2011; **24**: 3–10.
6. **Hald T, *et al.*** Development of a *Salmonella* source-attribution model for evaluating targets in the turkey meat production. Parma, Italy: European Food Safety Agency, 2012.
7. **European Food Safety Authority.** Overview of methods for source attribution for human illness from food borne microbiological hazards. *EFSA Journal* 2008; **764**: 1–43.
8. **Pires SM, *et al.*** Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathogens and Disease* 2009; **6**: 417–424.
9. **Hald T, *et al.*** A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Analysis* 2004; **24**: 255–269.
10. **van Pelt W, *et al.*** Origin, size and costs of human salmonellosis. Part 1. Origin of human salmonellosis related to pork, beef, chicken, eggs and other sources [in Dutch]. *Infectieziekten Bulletin* 1999; **10**: 240–243.
11. **Pires SM, Hald T.** Assessing the differences in public health impact of *Salmonella* subtypes using a Bayesian microbial subtyping approach for source attribution. *Foodborne Pathogens and Disease* 2010; **7**: 143–151.
12. **Wahlström H, *et al.*** Source attribution of human *Salmonella* cases in Sweden. *Epidemiology and Infection* 2011; **139**: 1246–1253.
13. **Guo C, *et al.*** Application of Bayesian techniques to model the burden of human salmonellosis attributable to U.S. food commodities at the point of processing: adaptation of a Danish model. *Foodborne Pathogens and Disease* 2011; **8**: 509–516.
14. **Pires S, *et al.*** Estimation of the relative contribution of different food and animal sources to human *Salmonella* infections in the European Union, Question No EFSA-Q-2010-00685. Parma, Italy: European Food Safety Agency, 2011.
15. **David JM, *et al.*** The Bayesian microbial subtyping attribution model: robustness to prior information and a proposition. *Risk Analysis*. Published online: 8 August 2012. doi:10.1111/j.1539-6924.2012.01877.x.
16. **Mullner P, *et al.*** Source attribution of food-borne zoonoses in New Zealand: a modified Hald model. *Risk Analysis* 2009; **29**: 970–984.
17. **Carrieri MP, *et al.*** Evaluation of the SIMI system, an experimental computerised network for the surveillance of communicable diseases in Italy. *European Journal of Epidemiology* 2000; **16**: 941–947.
18. **Busani L, *et al.*** Antibiotic resistance in *Salmonella enterica* serotypes Typhimurium, Enteritidis and Infantis from human infections, foodstuffs and farm animals in Italy. *Epidemiology and Infection* 2004; **132**: 245–251.
19. **Gelman A, Rubin DB.** Inference from iterative simulation using multiple sequences. *Statistical Science* 1992; **7**: 457–511.
20. **Hoelzer K, *et al.*** Animal contact as a source of human non-typhoidal salmonellosis. *Veterinary Research* 2011; **42**: 34.
21. **Collard JM, *et al.*** Drastic decrease of *Salmonella* Enteritidis isolated from humans in Belgium in 2005, shift in phage types and influence on foodborne outbreaks. *Epidemiology and Infection* 2008; **136**: 771–781.
22. **Marcus R, *et al.*** Dramatic decrease in the incidence of *Salmonella* serotype Enteritidis infections in 5 FoodNet sites: 1996–1999. *Clinical Infectious Diseases* 2004; **38**: S135–141.
23. **Cogan TA, Humphrey TJ.** The rise and fall of *Salmonella* Enteritidis in the UK. *Journal of Applied Microbiology* 2003; **94**: 114S–119S.
24. **Hendriksen RS, *et al.*** Global monitoring of *Salmonella* serovar distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: results of quality assured laboratories from 2001 to 2007. *Foodborne Pathogens and Disease* 2011; **8**: 887–900.

25. **Mullner P, *et al.*** Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. *Infection, Genetics and Evolution* 2009; **9**: 1311–1319.
26. **Franz E, van Bruggen AHC.** Ecology of *E. coli* O157:H7 and *Salmonella enterica* in the primary vegetable production chain. *Critical Reviews in Microbiology* 2008; **34**: 143–161.
27. **Podolak R, *et al.*** Sources and risk factors for contamination, survival, persistence, and heat resistance of *Salmonella* in low-moisture foods. *Journal of Food Protection* 2010; **73**: 1919–1936.