Epidemiol. Infect. (2015), **143**, 1148–1165. © Cambridge University Press 2014
doi:10.1017/S0950268814000429

Using surveillance and monitoring data of different origins in a *Salmonella* source attribution model: a European Union example with challenges and proposed solutions

L. V. DE KNEGT*, S. M. PIRES AND T. HALD

National Food Institute, Division of Epidemiology and Microbial Genomics, Technical University of Denmark, Denmark

Received 2 April 2013; Accepted 8 February 2014; first published online 15 July 2014

SUMMARY

Microbial subtyping approaches are commonly used for source attribution of human salmonellosis. Such methods require data on *Salmonella* in animals and humans, outbreaks, infection abroad and amounts of food available for consumption. A source attribution model was applied to 24 European countries, requiring special data management to produce a standardized dataset. Salmonellosis data on animals and humans were obtained from datasets provided by the European Food Safety Authority. The amount of food available for consumption was calculated based on production and trade data. Limitations included different types of underreporting, non-participation in prevalence studies, and non-availability of trade data. Cases without travel information were assumed to be domestic; non-subtyped human or animal records were re-identified according to proportions observed in reference datasets; missing trade information was estimated based on previous years. The resulting dataset included data on 24 serovars in humans, broilers, laying hens, pigs and turkeys in 24 countries.

Key words: Data management, food safety, *Salmonella*, source attribution, surveillance.

INTRODUCTION

Unsafe food is related to several kinds of diseases, ranging from diarrhoeal syndromes to various forms of cancer [1, 2]. In 2005, it was estimated that foodborne or waterborne diarrhoeal diseases were responsible for 2·2 million deaths per year worldwide, 1·9 million of which were children [2].

Salmonella spp. is one of the most common and widely distributed foodborne pathogens in the European Union (EU), with 108 614 laboratory-confirmed cases reported in 2009. Although its relative importance has been decreasing since 2006, *S.* Enteritidis is still the most frequently reported serovar (52·3% of cases), followed by *S.* Typhimurium (23·3%), and a wide range of other serovars of public health significance [3].

Identifying the main sources of an illness is a crucial step for the prioritization of control measures [4]. This process is called source attribution, and it can be achieved by a variety of methods, one of which is the microbial subtyping approach [5, 6]. The principle of this approach is to compare the occurrence of subtypes in animals or food sources with the same subtypes in humans, provided that subtypes are heterogeneously distributed among the sources. Human infections caused by source-specific subtypes

* Author for correspondence: Dr L. V. De Knegt, Technical University of Denmark, National Food Institute, Division of Epidemiology and Microbial Genomics, Mørkhøj Bygade 19, Building H, 2860 Søborg, Denmark.
(Email: ledkn@food.dtu.dk)

are attributed to the corresponding sources. Infections caused by subtypes found in several reservoirs are distributed relatively to the prevalence of the specific types.

The approach requires an integrated foodborne disease surveillance programme that collects isolates from the major food-animal reservoirs, sporadic human cases, outbreaks and travel-related cases [7]. For that reason, no EU-wide source attribution study has yet been conducted, as a unified European animal-and-human health database does not exist. However, since 2003, efforts have been made in the EU to standardize the reporting of pathogens and diseases in humans and animals, including the conduction of studies to estimate the Member State (MS)-level baseline prevalence of *Salmonella* in animals of the food chain [8–11]. Another initiative was the harmonization of the monitoring of *Salmonella* in laying hens [12], broilers [13] and turkeys [14], the last two implemented after the activities described in this paper were conducted. Data from harmonized animal programmes, other production animals and humans are summarized yearly in the European Union Summary Report on Trends and Zoonoses (EUSR) [3].

Based on a review of microbial subtyping-based source attribution studies [7, 15–17], the 'perfect' dataset for a EU-wide model would include, for each MS: (1) the number of reported salmonellosis cases in humans, originating from a nationally representative surveillance system in which cases are all laboratory-confirmed and subtyped to an appropriate discriminatory level; (2) information on whether the person had been travelling abroad 1 week prior to symptoms onset; (3) number of outbreak cases and identified outbreak sources; (4) for all major animal sources of human salmonellosis in Europe, the *Salmonella* prevalence using the same subtyping methods used for humans; and (5) the amount of an animal product originating from one country which is consumed in another country. Serotyping, combined with phage-type data and further differentiation based on antimicrobial resistance profiling, is currently considered the ideal level of subtyping for those models, as it better discriminates common subtypes (e.g. *S.* Enteritidis and *S.* Typhimurium) among similar sources, compared to using only serovars [18].

This paper describes the data obtained from sources available in 2010 to be used in an EU-wide source attribution microbial subtyping model and the data management steps taken to produce a sufficiently detailed and homogenous dataset containing *Salmonella* serovar information from humans and animal-food reservoirs. Limitations of the data available are presented, along with the solutions applied to solve them.

METHODS

Data sources

The European Surveillance System (TESSy). This is a system for collection, validation, analysis and dissemination of data from the EU and European Economic Area (EEA), administered by the European Centre for Disease Prevention and Control (ECDC) which has been functioning since 2008 [19]. Countries report their data on communicable diseases to the system, which also records information on outbreaks and the possibility of infection during international travel. Reporting of specific serovars is mandatory, but countries may report isolates as 'unknown', and further subtyping is only done on a voluntary basis. TESSy replaced the data collection systems for the Data Surveillance Network, which collected national data individually [19], therefore data from 2006 and 2007 on *Salmonella* exist in the system, but not in a completely standardized manner.

The European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Foodborne Outbreaks (EUSR). This report is prepared by the European Food Safety Authority (EFSA) and ECDC. Data on zoonoses and zoonotic agents in animals, foodstuffs and animal feed are reported annually by MS to EFSA and summarized in the EUSR. Serovar reporting follows the same requirements as for humans.

Baseline studies on the prevalence of Salmonella in animal populations in the European Union (BS). To provide the scientific basis for setting prevalence targets for reduction of *Salmonella* in commercial and breeding farms, EU-wide studies on the baseline prevalence of *Salmonella* were conducted on laying hen flocks (2004–2005) [8], broiler flocks (2005–2006) [11], slaughter pigs (2006–2007) [9], fattening and breeding turkeys (2006–2007) [10], broiler carcasses (2008) [20] and breeder pigs (2008) [21]. The studies took place in a 4-year period, and varied in participation due to new EU members in 2004 and 2007, and to the occasional participation of EEA countries. However, the studies still constitute the most uniformly collected

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 1. *Availability of data from the different datasets by country*

Source	Data source*	Country	Additional data sources
Laying hens	EUSR data 2008	AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, RO, SE, SI, SK, UK	
Cattle	EUSR data 2007–2009	AT, BE, BG, CH, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, RO, SE, SI, SK, UK	FR: David, 2009 [38] LV: EUSR 2006 [28]
Pigs	BS 2006, lymph node	AT, BE, BG, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, SE, SI, SK, UK	
Broiler	BS 2008, carcasses	AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, MT, NL, NO, PL, PT, RO, SE, SI, SK, UK	GR: BS 2005–2006 [11]
Turkey	BS 2006, fattening turkeys	AT, BE, BG, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, SE, SI, SK, UK	EE: EUSR 2006 [28] LU: EUSR 2008 [37] LV: EUSR 2006 [28]
Human cases	Foodborne outbreak data, 2007–2009	AT, BE, CH, CZ, DE, DK, EE, ES, FI, FR, HU, IE, LT, LV, NL, NO, PL, PT, RO, SE, SI, SK	
	TESSy case-based and aggregated data, 2007–2009†	AT, BE, BG, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, MT, NL, NO, PL, PT, RO, SE, SI, SK, UK	
	National monitoring and laboratory surveillance data 2007–2009‡	PL, PT, NL, IT, DE	

Austria (AT); Belgium (BE); Bulgaria (BG); Cyprus (CY); Czech Republic (CZ); Denmark (DK); Estonia (EE); Finland (FI); France (FR); Germany (DE); Greece (GR); Hungary (HU); Ireland (IE); Italy (IT); Latvia (LV); Lithuania (LT); Luxembourg (LU); Malta (MT); Norway (NO); Poland (PL); Portugal (PT); Romania (RO); Slovakia (SK); Slovenia (SI); Spain (ES); Sweden (SE); Switzerland (CH); The Netherlands (NL); United Kingdom (UK)

* If data were missing from a specific source in a country, used surrogate data sources are indicated.

† Bulgaria reported human cases, but no serovar information was available.

‡ Obtained through direct contact with Member States.

and analysed data on *Salmonella* at the EU-level, allowing valid comparisons between MS.

The Statistical Office of the European Union (EUROSTAT) [22]. EUROSTAT was established in 1953 to provide statistics at the European level to allow comparisons between countries and regions. It collects data on the value and quantity of food and slaughter animals traded among EU MS or with third-party countries. Although EU legislation ensures that the statistics provided to EUROSTAT are based on legal texts and harmonized definitions and procedures [22], information is provided directly by MS, and so data availability and quality are subject to variations in national focus and cultural differences.

Data collected from the sources described were stored and analysed in SAS Enterprise Guide, SAS/STAT® User's Guide, version 8 (SAS Institute Inc., USA).

Reported cases of human salmonellosis

Number and serovar distribution of sporadic human cases reported to TESSy from 2007 to 2009 were provided by ECDC through EFSA. Outbreak-related cases were provided by EFSA as reported by MSs. The total number of reported cases included sporadic, travel and outbreak-related infections. MS with a level of serovar detail insufficient for source attribution were requested to provide additional data, if available. Such national datasets were provided by Poland and Portugal. The MS providing data on sporadic and outbreak cases are summarized in Table 1.

Challenge 1: Underreporting

One issue arising from the use of surveillance data is the underreporting of cases. It is generally understood that the real number of illnesses occurring in the population is larger than the number that are reported to the surveillance system [23]. This is explained by the

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

percentage of: (1) people who seek medical care when sick; (2) people who provide clinical specimens when requested; (3) specimens which are tested; (4) sensitivity of the laboratory tests and (5) positive results that get reported [23]. Therefore, the true burden of human salmonellosis may be considerably larger than the officially reported incidence. The level of underreporting is expected to vary between countries, depending on differences in organization and effectiveness of local surveillance systems [19, 24].

Proposed solution. Havelaar *et al*. used data from a Swedish travel database and the *Salmonella* incidence from a Dutch population-based study to estimate a set of multipliers for correction of underreporting in 31 European countries [19, 24]. These underreporting factors (UF) were based on the proportion of cases of salmonellosis that were reported in Sweden upon returning from the Netherlands, and represent an estimation of the number of cases that should have been reported for each case that entered the system. The use of these multipliers is expected to have an impact on the most important sources estimated at EU-level. As the adjustment for underreporting is only done after the attribution process, the corrected numbers are not shown here, but can be found in de Knegt (2013) [18].

Challenge 2: Incomplete travel-related information

Travel information, derived from data reported as 'probable country of infection', was recorded as 'travel-related', 'domestic' or 'unknown'. The proportion of travellers and the amount of information provided is expected to vary among MS, depending on local habits and surveillance priorities.

Proposed solution. The Hald model and its adaptations [7, 15] use the observed proportion of travel cases that were properly discriminated to redistribute cases with no information to the 'travel-related' and 'domestic' categories, and the same approach could potentially be used in the EU model. In case there is not enough information available for redistribution, cases which did not specifically report a travel history should be considered domestic.

Challenge 3: Incomplete or missing serovar identification

Cases in which serovar identification is missing or incomplete can be summarized as: (*a*) classification up to genus or species level, e.g. *Salmonella* spp. or *Salmonella enterica*; (*b*) classification up to subspecies level, e.g. *Salmonella enterica enterica* or *Salmonella enterica* subspecies I; (*c*) classification using serogroups based on the O-antigen, e.g. B, C1–C2, O:4 or O:33; (*d*) main serovars properly specified and the rest aggregated as 'others'; (*e*) serovar field left blank or completed as 'unknown'.

Proposed solutions. Isolates not classified up to serovar level should be reassigned to specific serovars according to proportions observed in previous studies, in the same dataset or in other references, depending on the data availability in each case.

Isolates identified up to genus or species level, left blank or completed as 'unknown' should be reassigned to all serovars observed in the country (e.g. if *S.* Enteritidis accounts for 60% of all serotyped isolates in a country, and 10 isolates in the same country are not properly identified, six of them must be reassigned to *S.* Enteritidis). Isolates identified up to subspecies level should likewise be reassigned to all serovars in the country, but with proportions calculated using only isolates of *S. enterica enterica* as the total.

Isolates classified as serogroups should be distributed among serovars pertaining to those groups, in accordance with the Kauffman–White–Le Minor scheme [25] (e.g. if *S.* Typhimurium accounts for 40% of all isolates in a country, but for 80% of units from serovars belonging to group B, and 10 isolates are only identified as 'group B', eight of those should be reassigned to *S.* Typhimurium).

Isolates classified as 'others' are assumed to belong to serovars not described in the current dataset, but nonetheless present in the country. In this case, the reference used for reassignment of proportions is the World Health Organization Global Foodborne Infections Network (GFN) Country Databank (CDB) [26], which contains the 15 most commonly identified *Salmonella* serovars among human and non-human sources in 84 countries (e.g. a country reports 30 isolates to TESSy: 10 *S.* Enteritidis, 10 *S.* Typhimurium, and 10 'others'. The CDB shows 80% *S.* Enteritidis, 10% *S.* Typhimurium, 7% *S.* Infantis and 3% *S.* Hadar for this country, so, according to this reference, *S.* Infantis and *S.* Hadar correspond to 70% and 30% of the non-described serovars. The 10 isolates should be redistributed as seven *S.* Infantis and three *S.* Hadar, assuming that *S.* Typhimurium and *S.* Enteritidis are not included in the 'others' group).

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 2. *Order of priority for selection of animal-food data to include in the model*

Animal	First choice	Second choice	Third choice
Broilers	EU baseline survey (carcasses) (2008)	EU baseline survey (flocks) (2005–2006)	—
Pigs	EU baseline survey (2006–2007)	—	—
Turkeys	EU baseline survey (2006–2007)	EU surveillance and monitoring (2006 and 2008)	—
Layers	EU harmonized monitoring (2008)	EU baseline survey (2004–2005)	—
Cattle	EU surveillance and monitoring (herds) (2006–2009)	EU surveillance and monitoring (slaughter) (2007–2009)	Literature

Challenge 4: Underreporting and incomplete identification of serovars in outbreak data

For outbreaks of foodborne salmonellosis, the same datasets used for EUSRs 2007–2009 [3, 27, 28] were provided by EFSA. Not all countries report outbreak cases, and not all reported cases have complete serovar information.

Proposed solutions. The same underreporting multipliers used for sporadic cases cannot be applied to outbreaks, as cases belonging to an outbreak are likely to have a different probability of being reported, and some serovars may generate outbreaks more frequently than others [15]. Based on that, countries which report sporadic cases but no outbreak cases are assumed to have no foodborne *Salmonella* outbreaks in the period. Outbreak-related cases for which a serovar is not fully identified should be reassigned using the proportions observed in the same outbreak dataset.

Salmonella in livestock and food

Challenge 5: Heterogenous availability of data from MS and animal sources

The EU BS prevalence of *Salmonella* in the sources was the preferred data source. Due to the admission of new EU members and the voluntary character of BS participation, data were not available for all MS and sources. However, these datasets were considered the most representative of the given reservoirs, since no harmonized EU monitoring in pigs and turkeys was currently in place. In addition, the broiler carcass study was considered to provide more recent data than BS on broiler flocks, and with a better detailing of the serovar distribution compared to the existing EU monitoring data. The laying-hen BS was conducted between 2004 and 2005 [8], and it is expected that the implementation of harmonized monitoring [12] has resulted in significant changes in *Salmonella* serovar prevalences in this reservoir in many MS. No data from BS or EU-harmonized monitoring exist for cattle.

Proposed solutions. In order to use the most recent data possible, data that are missing from BS should be supplied with surveillance and monitoring data found in the EUSR. When not enough surveillance or monitoring data at the herd/flock level are available for a source or MS, slaughter samples should be surveyed and their quality as substitutes assessed. The order of priority for selecting which animal-food data to include in the model is shown in Table 2.

Challenge 6: Incomplete or missing serovar identification

The expected situations in which serovar identification is missing or incomplete are the same as for human data. For BS data, no reference for reassigning serogroups or incomplete serovar identification was available.

Proposed solutions. The criteria for reassigning non-identified or partially identified serovars should be the same as for the human data. Proportions found in the laying-hen BS [8] should be used for reallocation of laying-hen monitoring data. In datasets where there are no records identified as 'others', units should be redistributed according to the proportions found in properly identified serovars in the same dataset.

Food production and trade data

Food production data were derived by EFSA from the EUROSTAT databases on production and slaughtered animals for food consumption [22]. Consumption calculations were based on production and country-to-country trade data. This was done so the

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

attribution model can account for the amount of food present in a given country that originated from other countries, and use the country- and food-specific serovar prevalences for the attribution [18]. The domestically produced amount available for consumption of a food source in a MS was estimated as domestic production minus export, whereas the amount of imported food available for consumption in MS A originating from MS B was estimated as import minus re-export (when relevant). For this study, extra EU food trade was not considered [18].

Challenge 7: Missing data

Information on poultry for meat production was not available for Belgium in 2007 and 2008. Egg production data was lacking for several countries, and data for most food sources and most years were missing for Cyprus. Data on the export of the food sources to other MS included in this study were available for all considered countries, with the exception of the amount of eggs exported from Cyprus.

Proposed solutions. Missing data on annual quantities of poultry meat products sold per MS, with differentiation between boilers, turkeys and other poultry species are available in the 2009 Annual Report of the Association of Poultry Processors and Poultry Trade in the EU Countries (AVEC) [29]. For all sources, countries with information missing for a year should have the missing value estimated based on the percentage of increase or decrease between available years; when data from only one year are available, that value should be used as surrogate for the missing years.

Challenge 8: Negative estimated amounts available for consumption

Due to differences in numbers reported in the production, import and export datasets, the calculations of the amount of a food source available for consumption in a country in some cases results in negative numbers, meaning that the volume exported is larger than the domestic production.

Proposed solution. In order to ensure that MS will still have nationally produced food available in their own country, re-exporting of imported products should be considered possible.

Challenge 9: Validation of the estimation of consumption data based on trade data

The underlying assumptions for this estimation were that EUROSTAT data were complete and consistent, and that all food available for consumption is actually consumed, in a way that these data reflect the real flow of foodstuffs and consequent exposure in the countries. According to an assessment performed by EFSA [30], the information recorded in those datasets does not fully support these assumptions. The assessment showed the existence and non-reporting of triangular trade, misclassification of food products and problems in the conversion of currency/weight units. Moreover, we expect that in several situations, data for missing years needs to be estimated or supplied with surrogate data (e.g. AVEC data), resulting in a highly manipulated dataset that may not represent reality.

Proposed solution. The data management can be validated by comparing the resulting consumption dataset with consumption data available from the WHO Global Environment Monitoring System Food Consumption Cluster Diets [31]. As the WHO data only offer the broad category 'poultry', broilers and turkeys should be summed. Relative proportions of consumption of poultry, pork and eggs must be calculated, so a proportional similarity index (PSI/Czekanowsky index) can be used to compare those proportions between the two groups in each country. The PSI is an estimate of the area of intersection between two frequency distributions [32], calculated as

$$\mathrm{PSI} = 1 - 0{\cdot}5 * \sum |p_1 - q_1| = \sum \mathrm{Min}\,(p_1, q_1).$$

It is traditionally used for calculating niche overlap and resource availability in population ecology [33] or proportions of identified bacterial strains in epidemiology [34, 35], but it was considered that each of the relative proportions in the three sources corresponds to the area under a probability curve, and so the same measure could be applied. A PSI of 1 means a complete overlap, or 100% similarity. An 'overall PSI' for the whole dataset was calculated by averaging the country PSI values.

RESULTS

Human data

The percentage of records with incomplete identification and that had to be reassigned varied from zero in Portugal to 84% in Romania (Table 3). The most

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 3. *Number and percentage of reassigned records in humans*

	Incomplete identification												
	Species/genus*		Subspecies†		Serogroup‡		Aggregated data§		Unknown¶		Total reported	Reassigned	
Country	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%		*n*	%
AT			2	0·02	132	1·56	287	3·38	362	4·27	8487	783	9·23
BE							172	1·55			11066	172	1·55
BG	—	—	—	—	—	—	—	—	—	—	3899	—	—
CY	2	0·42			9	1·91			101	21·44	471	112	23·78
CZ									586	1·51	38842	586	1·51
DE			462	0·36	8057	6·33	5782	4·54	1628	1·28	127330	15929	12·51
DK			2	0·03	3	0·04	25	0·33	342	4·56	7497	372	4·96
EE					25	1·86	28	2·09			1341	53	3·95
ES							2504	20·81	2,091	17·38	12033	4595	38·19
FI	19	0·23	3	0·04	23	0·28	6	0·07	22	0·27	8228	73	0·89
FR							2185	10·75			20319	2185	10·75
GR					104	5·40	3	0·16	1309	67·93	1927	1416	73·48
HU			57	0·30	191	1·00	908	4·76	2	0·01	19091	1158	6·07
IE	1	0·08					11	0·87	68	5·38	1264	83	6·57
IT	25	0·24			6	0·06			1080	10·58	10205	1111	10·89
LT					56	0·73	156	2·04	191	2·50	7643	403	5·27
LU									63	13·15	479	63	13·15
LV							53	1·99	608	22·81	2665	661	24·80
MT	20	5·39							40	10·78	371	60	16·17
NL			210	5·04			84	2·02			4168	294	7·05
PL							1204	3·89			30963	1204	3·89
PT											1513	0	0·00
RO							1218	51·81	766	32·58	2351	1984	84·39
SE			68	0·60			411	3·65	307	2·73	11265	786	6·98
SI					63	2·10					3002	63	2·10
SK	3	0·02			154	0·79	84	0·43	87	0·45	19399	328	1·69
UK	7	0·02			149	0·41	4	0·01	1009	2·75	36666	1169	3·19
EU total	77	0·02	804	0·20	8975	2·29	15125	3·85	10662	2·72	392485	35643	9·08
CH	—	—	—	—	—	—	—	—	—	—	—	—	—
NO							21	0·44	10	0·21	4825	31	0·64
Total	77	0·02	804	0·20	8975	2·26	15146	3·81	10672	2·69	397310	35674	8·98

For explanation of country abbreviations see Table 1.

* *Salmonella* spp, *Salmonella enterica, Salmonella* not typed, *Salmonella* untyped.

† *Salmonella enterica enterica*, subspecies I.

‡ B, C, D, E, D1, C1, C2–C3, D1, E1.

§ 'Others', 'other serovars'.

¶ 'Unknown'.

Table 4. *Number and percentage of reassigned records in foodborne Salmonella outbreaks*

		Incomplete identification						
		Species/genus*		Serogroup†			Reassigned	
Country	Reported	*n*	%	*n*	%	Total reported	*n*	%
AT	Yes					421	0	0·00
BE	Yes					91	0	0·00
BG	No					—	—	—
CY	No					0	0	0·00
CZ	Yes					337	0	0·00
DE	Yes	13	0·55			2383	13	0·55
DK	Yes					2224	0	0·00
EE	Yes					157	0	0·00
ES	Yes					469	0	0·00
FI	Yes					189	0	0·00
FR	Yes	1218	46·68			2609	1218	46·68
GR	No					0	0	0·00
HU	Yes	86	4·48			1921	86	4·48
IE	Yes					67	0	0·00
IT	No					0	0	0·00
LT	Yes					371	0	0·00
LU	No					0	0	0·00
LV	Yes	201	39·26			512	201	39·26
MT	No					0	0	0·00
NL	Yes	12	1·71	26	3·71	700	38	5·43
PL	Yes			29	0·55	5310	29	0·55
PT	Yes					90	0	0·00
RO	Yes	26	5·95			437	26	5·95
SE	Yes	8	2·94			272	8	2·94
SI	Yes					692	0	0·00
SK	Yes					583	0	0·00
UK	No					0	0	0·00
EU total	—	1564	7·89	55	0·28	19835	1619	8·16
CH	Yes					6	0	0·00
NO	Yes					95	0	0·00
Total	—	1564	7·85	55	0·28	19936	1619	8·12

For explanation of country abbreviations see Table 1.
* *Salmonella enterica enterica*, subspecies I.
† B, C, D, E, D1, C1, C2–C3, D1, E1.

common reason for reassignment was records reported in aggregated form, i.e. with several serovars categorized as 'others', and the next reason was isolates reported as 'unknown', followed by isolates only classified as serogroup (Table 3). Besides the predicted identification problems, a specific issue regarding *S*. Typhimurium was found: one of the defining characteristics of *S*. Typhimurium is the two phases of the H-antigens: 'i' and '1,2', which is why the antigenic formula for this serovar is written as '1,4,[5],12:i:1,2' [27]. However, *S*. Typhimurium-like variants with only the first phase of the H-antigen (e.g. 1,4,[5],12:i:- or 1,4,[5],12:-:-) have been reported, and are referred to as '*S*. Typhimurium-like strains' or 'monophasic *S*. Typhimurium' [36]. For our purposes, those isolates were reassigned to *S*. Typhimurium, which is supported by an EFSA Biohazard Panel assessment [36].

Bulgaria, Cyprus, Greece, Italy, Luxembourg, Malta and the UK did not report outbreak cases. Nearly 47% of outbreak cases reported by France had to be reassigned, as the isolates were reported as '*Salmonella* spp'. For Latvia, the proportion was 39% (Table 4). Switzerland reported outbreaks, but no sporadic cases (Table 1).

Travel information (Table 5) was reported as 'unknown' for 100% of isolates in France, Romania

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 5. *Number of cases reported in the original datasets as travel-related, domestic or unknown and the total used in the model, assuming that any case not specifically mentioned as travel-related was domestic*

	Reported			Total used	
Country	Travel	Domestic	Unknown	Travel	Domestic
AT	988	7499	0	988	7499
BE	0	11066	0	0	11066
BG	—	—	—	—	—
CY	18	428	25	18	453
CZ	657	38185	0	657	38185
DE	6683	114362	6285	6683	120647
DK	1366	2645	3486	1,366	6131
EE	95	1246	0	95	1246
ES	0	12033	0	0	12033
FI	6845	1059	324	6845	1383
FR	0	0	20319	0	20319
GR	45	1763	119	45	1882
HU	29	19062	0	29	19062
IE	384	343	537	384	880
IT	132	692	9381	132	10073
LT	21	0	7622	21	7622
LU	46	431	2	46	433
LV	32	1817	816	32	2633
MT	4	365	2	4	367
NL	497	3671	0	497	3671
PL	16	0	30947	16	30947
PT	5	0	1508	5	1508
RO	0	0	2351	0	2351
SE	8752	2207	306	8752	2513
SI	0	0	3002	0	3002
SK	146	19253	0	146	19253
UK	8921	8084	19661	8921	27745
EU total	35682	246211	106693	35682	356803
CH	—	—	—	—	—
NO	3721	871	233	3721	1104
Total	39403	247082	106926	39403	357907

For explanation of country abbreviations see Table 1.

and Slovenia. Full travel information was provided by Austria, Belgium, the Czech Republic, Estonia, Spain, Hungary, The Netherlands and Slovakia. The remaining MS had variable proportions of cases reported as 'travel-related', 'domestic' or 'unknown'. Therefore, the proposed 'informed redistribution' was not possible, as a large number of countries did not report any travel cases. As a consequence, all records with missing or unknown travel information from countries with serovar details of sporadic cases were considered domestically acquired in the reporting country.

Table 6 shows the relative occurrence of the 11 most important zoonotic serovars in the last 5 years in sporadic and outbreak cases [3, 28]. *S.* Enteritidis and *S.* Typhimurium were the most frequently observed in sporadic cases, along with *S.* Infantis, *S.* Newport, *S.* Kentucky, *S.* Virchow, *S.* Derby and *S.* Agona. The most commonly observed serovars in outbreaks were also *S.* Enteritidis and *S.* Typhimurium. As expected, outbreaks may present serovars not normally found in a specific country. That is particularly true in countries with a small number of sporadic cases and good *Salmonella* control in domestic products, e.g. Finland or Sweden.

Animal-food data

Data was available from 28 countries (Table 1). Laying-hen data from the EUSR 2008 [28] were preferred over BS data, as this was the first year of EU-harmonized reporting for this reservoir, Greece did not participate in the broiler carcasses study [20],

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 6. *Total isolates and relative proportions of the most frequent serovars in total reported (R) and outbreak (O) cases in humans in the EU and Norway, 2007–2009*

Country	Reporting†	Serovar* 1	2	3	4	5	6	7	8	9	10	11	12	Total isolates
AT	S	71·9	11·3	1·3	0·3	0·7	0·4	0·7	0·1	0·4	0·3	0·2	12·4	8487
	O	51·5	46·1	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	2·4	421
BE	S	21·6	57·6	1·0	0·8	0·7	0·3	0·7	1·4	0·7	0·2	0·4	14·7	11 066
	O	57·1	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	42·9	91
CY	S	45·2	13·0	0·8	0·2	2·3	0·8	0·6	0·8	1·7	1·9	0·0	32·5	471
	O‡	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CZ	S	92·2	4·8	0·6	0·1	0·2	0·1	0·2	0·1	0·2	0·0	0·1	1·4	38 215
	O	100·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	337
DE	S	58·6	27·9	1·4	0·2	0·4	0·2	0·2	0·4	0·3	0·1	0·4	9·9	127 330
	O	87·7	2·8	2·2	0·0	0·1	0·0	0·0	0·0	0·0	0·0	2·6	4·6	2383
DK	S	25·5	44·0	1·2	0·6	1·6	2·1	0·3	1·1	1·8	1·7	0·9	19·3	7497
	O	13·9	84·0	0·0	0·0	0·0	0·0	0·0	0·4	0·0	0·0	0·0	1·6	2224
EE	S	82·6	8·8	1·2	0·4	0·4	0·2	0·1	0·4	0·1	0·3	0·4	5·0	1341
	O	96·2	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	3·8	0·0	157
ES	S	47·1	26·6	0·2	0·0	0·1	0·0	0·2	3·5	7·7	0·0	0·0	14·7	12 033
	O	47·5	14·9	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	37·5	469
FI	S	35·8	12·8	1·9	0·9	4·8	1·9	0·9	0·5	3·9	6·5	1·3	28·7	8228
	O	0·0	9·5	0·0	0·0	0·0	21·2	0·0	0·0	50·3	0·0	14·8	4·2	189
FR	S	17·8	31·5	1·5	1·8	1·1	0·8	1·9	2·2	1·8	0·3	0·7	38·6	20 319
	O	33·6	50·1	0·0	0·0	1·2	0·0	0·1	0·0	8·5	0·0	2·1	4·3	2609
GR	S	80·6	7·5	0·6	0·0	1·2	0·0	0·1	0·0	0·1	0·0	0·3	9·7	1927
	O‡	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
HU	S	70·6	12·6	5·4	0·0	0·3	0·1	0·4	0·5	0·1	0·1	0·7	9·2	19 091
	O	96·7	1·4	1·6	0·0	0·0	0·0	0·0	0·0	0·1	0·0	0·0	0·3	1921
IE	S	32·8	30·1	1·2	1·7	1·7	1·9	0·7	0·4	1·9	0·7	0·1	26·9	1264
	O	77·6	6·0	0·0	0·0	0·0	16·4	0·0	0·0	0·0	0·0	0·0	0·0	67
IT	S	19·1	50·8	2·1	0·2	0·5	0·3	1·1	2·9	0·6	0·1	0·3	22·1	10 205
	O‡	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
LT	S	91·2	5·3	0·6	0·0	0·1	0·3	0·0	0·4	0·1	0·1	0·0	1·8	7643
	O	100·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	371
LU	S	48·0	20·0	0·4	0·6	0·8	0·4	0·4	2·3	0·8	0·2	0·4	25·5	479
	O‡	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Table 6 (*cont.*)

Country	Reporting†	Serovar*												Total isolates
		1	2	3	4	5	6	7	8	9	10	11	12	
LV	S	90·8	5·0	0·0	0·0	0·8	0·1	0·1	0·0	0·2	0·3	0·8	2·0	2665
	O	97·5	1·4	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	1·2	512
MT	S	54·3	27·1	3·0	1·2	0·3	0·0	0·0	1·5	0·0	0·0	0·0	12·5	371
	O‡	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NL	S	36·2	33·3	1·5	1·3	1·5	0·4	0·4	0·5	1·6	0·6	0·4	22·3	4168
	O	35·7	59·4	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	4·9	700
NO	S	46·5	17·9	1·2	0·5	2·5	1·7	0·6	0·4	1·7	4·4	0·3	22·2	4825
	O	0·0	61·1	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	38·9	95
PL	S	80·5	8·9	2·0	0·0	2·9	0·3	1·5	0·1	0·4	0·0	0·0	3·4	30 963
	O	93·1	2·0	2·0	0·0	1·1	0·0	1·0	0·0	0·2	0·0	0·0	0·6	5310
PT	S	61·7	23·1	0·3	0·0	0·0	0·1	0·1	0·5	0·5	0·0	0·6	13·0	1513
	O	95·6	4·4	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	90
RO	S	39·6	40·5	1·2	0·0	1·8	1·2	0·0	1·5	0·6	0·0	0·0	13·8	1585
	O	51·5	29·1	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	19·5	437
SE	S	39·1	11·4	0·7	0·5	3·5	0·9	0·6	0·3	1·6	5·6	0·3	35·7	11 265
	O	0·0	1·8	0·0	0·0	0·0	0·0	0·0	0·0	0·0	19·1	0·0	79·0	272
SI	S	79·7	5·6	1·1	0·1	0·2	0·2	0·1	0·1	0·1	0·1	0·0	12·7	3002
	O	99·6	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·4	692
SK	S	87·8	6·7	1·5	0·1	0·3	0·2	0·3	0·1	0·1	0·1	0·6	2·3	19 399
	O	97·9	2·1	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	583
UK	S	45·9	17·5	1·3	1·2	2·9	1·3	0·8	0·3	1·8	1·2	0·2	25·5	36 666
	O‡	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

For explanation of country abbreviations see Table 1.

* 1, *S.* Enteritidis; 2, *S.* Typhimurium; 3, *S.* Infantis; 4, *S.* Kentucky; 5, *S.* Virchow; 6, *S.* Agona; 7, *S.* Hadar; 8, *S.* Derby; 9, *S.* Newport; 10, *S.* Stanley; 11, *S.* Bovismorbificans; 12, others; n.a., not available.

† S = sporadic cases; O = outbreak cases.

‡ The country did not report any outbreak cases.

Table 7. *Number and percentage of records reassigned to serovars in animal reservoirs*

	Incomplete identification										
	Species/genus*		Subspecies†		Serogroup‡		Aggregated§		Total positives	Reassigned	
Country	*N*	%	*n*	%	*n*	%	*n*	%		*N*	%
Broilers											
BE	15	19·48							77	15	19·48
IT	13	19·70							66	13	19·70
LT	15	57·69							26	15	57·69
MT	10	12·99							77	10	12·99
NL	1	2·33							43	1	2·33
Pigs											
BG			4	11·43					35	4	11·43
CY	5	10·64	3	6·38	1	2·13			47	9	19·15
DE	5	1·54			64	19·69			325	69	21·23
EE			4	14·81					27	4	14·81
ES	62	7·69							806	62	7·69
FR	5	2·33							215	5	2·33
GR	3	4·11	8	10·96					73	11	15·07
IE	1	1·54							65	1	1·54
IT	41	35·34	6	5·17					116	47	40·52
LV	2	9·52							21	2	9·52
NL	2	2·17	2	2·17					92	4	4·35
SI	4	14·81							27	4	14·81
Turkeys											
CY					5	17·86			28	5	17·86
DE					11	10·19			108	11	10·19
DK	1	100·00							1	1	100·00
HU	1	0·11	2	0·22					915	3	0·33
IT			8	2·89					277	8	2·89
SI					1	1·00			100	1	1·00
Layers											
AT	2	4·08							49	2	4·08
BE	3	3·95			3	3·95			76	6	7·89
CY					1	20·00			5	1	20·00
DE	13	5·91					23	10·45	220	36	16·36
ES	186	49·47							376	186	49·47
FR	20	10·70					6	3·21	187	26	13·90
HU							26	25·74	101	26	25·74
IT							115	67·25	171	115	67·25
PL							29	15·10	192	29	15·10
PT							9	10·84	83	9	10·84
UK							16	23·88	67	16	23·88
Bovines											
BE	3	3·70			4	4·94			81	7	8·64
DE	4	2·45					36	22·09	163	40	24·54
DK	4	44·44							9	4	44·44
ES	13	44·83							29	13	44·83
HU	25	80·65							31	25	80·65
IT	4	23·53							17	4	23·53
LU	1	14·29							7	1	14·29
NL	1	5·56							18	1	5·56
SE	6	10·00							60	6	10·00
UK	824	92·07							895	824	92·07

For explanation of country abbreviations see Table 1.

* *Salmonella* spp., *Salmonella enterica, Salmonella* not typed, *Salmonella* untyped.

† *Salmonella enterica enterica*, subspecies I.

‡ B, C, D, E, D1, C1, C2–C3, D1, E1.

§ 'Others', 'other serovars'.

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 8. *Relative proportions of the top-10 Salmonella serovars found in broiler carcasses, pig lymph nodes, turkey flocks and laying hen flocks in the chosen datasets*

Serovar*	Broilers†	Pigs‡	Turkeys§	Layers
S. Infantis	**29·2**	**1·9**	**6·6**	**11·5**
S. Enteritidis	**13·6**	**4·9**	**5·1**	**59·9**
S. Kentucky	**6·2**	0·0	0·1	0·0
S. Typhimurium	**4·4**	**44·9**	**7·9**	**8·3**
S. Bredeney	**4·3**	**2·0**	**17·2**	1·0
S. Virchow	**4·1**	0·3	1·0	**2·7**
S. Hadar	**3·8**	0·3	**14·0**	**3·4**
S. Paratyphi var. Java	**3·8**	0·1	0·2	0·1
S. Agona	**3·0**	**1·1**	2·9	**2·2**
S. Indiana	**2·9**	0·1	3·0	0·3
S. Derby	0·8	**14·6**	**11·3**	0·0
S. Rissen	0·0	**5·8**	0·0	0·5
S. Anatum	0·7	**2·4**	0·4	0·7
S. London	0·0	**1·3**	2·9	0·0
S. Brandenburg	0·2	**1·2**	0·0	0·9
S. Saintpaul	0·2	0·1	**10·3**	0·0
S. Kottbus	0·7	0·3	**8·3**	0·0
S. Orion	0·0	0·0	**6·1**	0·0
S. Blockley	1·8	0·1	**3·7**	0·0
S. Mbandaka	2·4	0·3	0·8	**6·6**
S. Livingstone	1·0	0·4	0·0	**3·4**
S. Ohio	0·9	0·3	0·0	**2·4**
S. Braenderup	0·2	0·2	0·1	**2·0**

* Combined list of the top ten serovars in all BS. Bold values show the top ten serovars for each animal reservoir.
† [20]. Participating countries: AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, MT, NL, NO, PL, PT, RO, SE, SI, SK, UK.
‡ [9]. Participating countries: AT, BE, BG, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, SE, SI, SK, UK.
§ [10]. Participating countries: AT, BE, BG, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, SE, SI, SK, UK.
¶ [27] Participating countries: AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GR, HU, IE, IT, LT, LU, LV, NL, NO, PL, PT, RO, SE, SI, SK, UK.

being supplied with data from the broiler flocks BS [11]. Malta and Romania did not participate in the slaughter pigs BS [9], and no surrogate data was available for those countries. For turkeys, BS data from fattening flocks were chosen [10], with the exception of Estonia, Latvia, Luxembourg and Romania, which were not part of the study. Data for Estonia, Latvia and Luxembourg were retrieved from EUSR 2006 and 2008 [28, 37]. No surrogate data was available for Romania. Non-harmonized surveillance data on cattle, including carcass samples at slaughter, were retrieved from EUSR 2007, 2008 and 2009 [3, 27, 28], with 2009 data being preferred to the other years. Cattle data for France was retrieved from a PhD thesis [38]. For this reservoir, no data from Cyprus or Malta were identified, and for some countries only one year of data was available. Belgium and the UK only reported positive samples for cattle, resulting in 100% positivity in those countries. Small samples were observed for broilers in Luxembourg, laying hens in Lithuania and Luxembourg and turkeys in Estonia, Luxembourg and Latvia. The amount and percentage of reassigned records in the total positives are given in Table 7.

Serovar predominance varied between countries in all animal sources. Considering the relative occurrence of serovars and number of countries in which they predominated, *S.* Infantis and *S.* Enteritidis were the main serovars observed in broilers, *S.* Typhimurium and *S.* Derby in pigs, *S.* Typhimurium, *S.* Bredeney and *S.* Hadar in turkeys and *S.* Enteritidis and *S.* Infantis in layers. *S.* Dublin and *S.* Typhimurium were the main serovars in cattle, but the data was considered too heterogeneous and frail to be representative. The top-ten serovars for broilers, pigs, turkeys and layers are given in Table 8.

Trade and consumption data

Availability of data on the annual quantities of poultry, pork, bovine meat and eggs produced varied by year and MS. All MS reported imports from other MS for all food products in the study period. The resulting surrogate consumption dataset was considered valid, as shown by the results of the data validation (Table 9). The individual PSI values were $>0{\cdot}8$ in most countries, indicating more than 80% similarity between the estimated data and the reference values. The one exception was Cyprus, with only 42% similarity, which is expected to have an impact on the attribution estimates for this country. The overall PSI was 0·91, indicating that the dataset as a whole can be used without considerable bias.

Final dataset for the source attribution model

Based on data availability and quality, 24 countries were included in the model: Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, The Netherlands and UK. Countries initially analysed and excluded were Bulgaria, which provided 100% of human cases

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

Table 9. *Comparison of the relative proportion of consumption of pork, poultry meat and table eggs in the WHO GEMS/Food data and the surrogate values calculated from EUROSTAT data*

	WHO GEMS/Food (%)			EUROSTAT (%)			
Country	Poultry	Pig	Egg	Poultry	Pig	Egg	PSI
AT	16·7	70·9	12·4	18·8	68·8	12·4	0·98
BE	32·3	50·5	17·2	28·7	58·1	13·2	0·92
CY	38·7	48·3	13·0	96·8	2·9	0·3	0·42
CZ	28·6	52·7	18·6	28·4	52·9	18·7	1·00
DE	17·4	67·0	15·6	24·1	63·2	12·7	0·93
DK	19·4	64·2	16·5	13·1	81·3	5·6	0·83
EE	33·5	47·6	18·8	33·4	49·7	16·9	0·98
ES	25·8	61·0	13·2	30·9	56·2	12·9	0·95
FI	25·8	58·7	15·5	24·5	49·9	25·6	0·90
FR	32·9	47·7	19·4	42·1	39·5	18·4	0·91
GR	31·5	53·1	15·4	33·2	47·9	18·9	0·95
HU	33·2	49·8	17·0	41·0	42·0	17·1	0·92
IE	36·3	54·7	9·0	40·9	45·7	13·4	0·91
IT	24·4	59·9	15·7	31·0	53·9	15·1	0·93
LT	24·6	51·4	23·9	30·7	51·1	18·2	0·94
LU	47·8	44·3	8·0	32·2	45·7	22·1	0·84
LV	30·3	44·7	25·0	33·6	43·0	23·4	0·97
NL	16·2	59·6	24·2	31·0	51·5	17·5	0·85
PL	23·8	61·7	14·5	31·3	56·6	12·0	0·92
PT	32·7	54·2	13·1	34·8	50·7	14·5	0·97
SE	20·9	61·3	17·8	22·3	58·6	19·1	0·97
SI	37·9	50·9	11·2	44·6	39·2	16·2	0·88
SK	36·5	45·8	17·7	28·2	48·7	23·1	0·92
UK	44·2	38·7	17·1	48·0	33·7	18·3	0·95
Overall PSI							0·91

For explanation of country abbreviations see Table 1.
PSI, Proportional similarity index.

without serovar details; Romania, which only participated in one BS, did not have enough surrogate data to be retrieved from the EUSR, and reported 84% of human cases without serovar information; Norway and Switzerland, which do not report to EUROSTAT, the latter also does not report to TESSy.

Based on the availability of EU-wide homogeneous data or good-quality surrogates, food-animal sources included were broilers, pigs, turkeys and laying hens (as the animal reservoirs for chicken meat, pork, turkey meat and eggs). Due to better completeness and availability, the resulting trade data from 2009 was used as consumption data for those sources. Cattle data were in general poor, and for some MS consisted of clinical isolates only. The use of herd information from 2007–2008 or slaughterhouse carcass samples was not sufficient to obtain a representative dataset for this source.

Twenty-two serovars were selected to be specifically addressed, based on their presence and importance in humans and chosen animal reservoirs: *S.* Agona, *S.* Anatum, *S.* Bovismorbificans, *S.* Braenderup, *S.* Brandenburg, *S.* Bredeney, *S.* Derby, *S.* Enteritidis, *S.* Hadar, *S.* Heidelberg, *S.* Infantis, *S.* Kentucky, *S.* Kottbus, *S.* Livingstone, *S.* London, *S.* Mbandaka, *S.* Montevideo, *S.* Newport, *S.* Rissen, *S.* Saintpaul, *S.* Typhimurium and *S.* Virchow. Albeit important in humans in most of the 24 countries, *S.* Stanley was not isolated from any of the selected reservoirs, while *S.* Dublin and *S.* Ohio became irrelevant after cattle were excluded. The building of the final *Salmonella* dataset (trade data not included) is shown in Figure 1.

DISCUSSION

This study presented the officially reported data available for use in an EU *Salmonella* source attribution model based on microbial subtyping [18]. Challenges

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

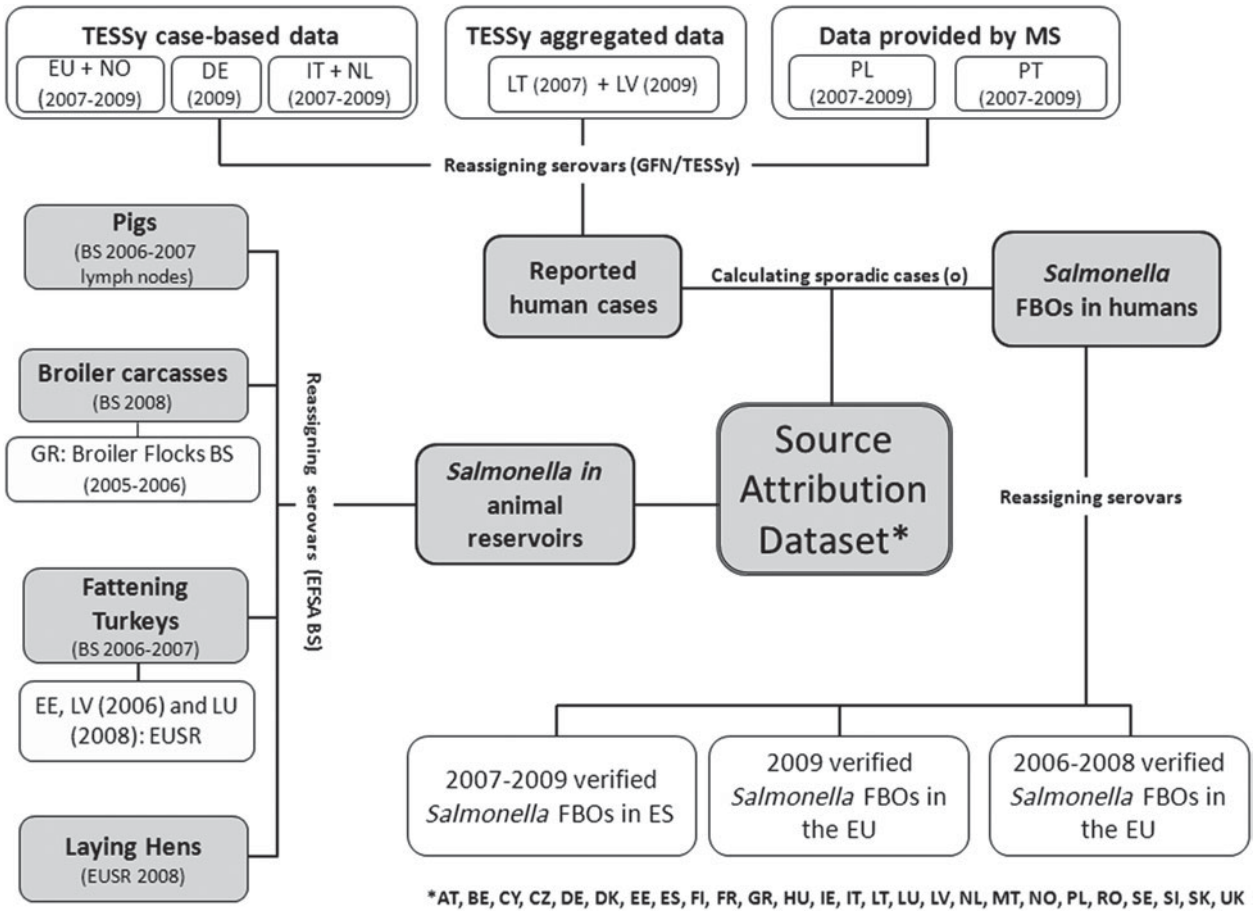

Fig. 1. The final *Salmonella* dataset (not including trade data). * For abbreviations see Table 1. FBO, Foodborne outbreaks.

associated with the use of these data were presented, and solutions were proposed. The data available were retrieved from multiple sources and had varied levels of quality and completeness. Although TESSy and EFSA collect and organize the data at the EU-level in a harmonized way, the primary information is collected in different countries, which have their own individual approaches and methods for data collection and management. Non-EU countries such as Switzerland and Norway also contribute to data heterogeneity, as they participate in some studies and report partial data, for example, to the EUROSTAT production database, but not to the trade database. This variability made several data management steps necessary.

The variability observed in the number of reported human *Salmonella* infections reflects true differences in the burden of salmonellosis across countries, but also differences in foodborne disease surveillance systems in MS and different levels of underreporting. The loss of data at various points along the surveillance chain from patient to official statistics is recognized in all countries [23], and multiplying factors [24] were used to compensate this loss. Limitations and assumptions connected to the use of those factors should be discussed, as they were calculated based on Swedish cases [24], coming from a system where underreporting is also expected to occur. By using the infection rates in returning travellers to calculate incidences for the local population in the countries visited, it was assumed that the eating habits and other exposures of Swedish travellers are the same as for the locals, also disregarding local levels of acquired immunity and differences in circulating strains. Considerations must also be made regarding the use of a Dutch study as a reference to estimate the underreporting in other countries, and a full discussion of the limitations can be found in Havelaar *et al.* [24]. Despite these limitations, the UF-adjusted numbers are still a better reflection of reality than the raw reported data, and this adjustment is expected to affect the relative importance attributed to the different sources by the model at the EU level, as it affects the

https://doi.org/10.1017/S0950268814000429 Published online by Cambridge University Press

contribution of each country to the total burden of salmonellosis in the EU.

Information about travelling within or outside Europe was not available in a representative manner, and it was not possible to estimate additional 'extra' intra-EU travellers because the proportion of reported cases with missing travel information was 100% in some countries. Thus, it had to be assumed that all reported cases with missing travel information were domestically acquired, which is expected to overestimate domestic cases, since travel information as reported to TESSy is often incomplete and may not reflect the true relationship between travel and domestic cases [3].

Concerning animal data, the panel of participating MS varied with each BS, as countries have the right to refuse participation in EU-wide baseline studies. The admittance of new MS to the EU also generates different lists of reporting countries for each animal source, as data were collected in different years. The resulting data gaps were, when possible, filled with information from EUSR. There are currently no EU-wide studies on the baseline prevalence of *Salmonella* in cattle and no harmonized monitoring in place, which is the main reason why this reservoir was excluded. However, this is not expected to have a large impact on the model, as national attribution studies have suggested that the contribution from the cattle reservoir in general is low compared to the other sources [15].

Serovar information was also heterogeneous both in humans and animals. Countries were approached directly for more complete datasets, and records were reassigned based on the serovar distributions observed in available data or external reference datasets (e.g. WHO GFN/CDB). One limitation of this approach is that any emergence of new serovars or other profile fluctuations may be lost, particularly when a whole year of typing is missing and the records are reassigned based on data from previous years. This is also a problem for outbreak cases, as two MS had nearly 50% of reassigned records, while others had the reference proportions calculated from a small number of reported cases. Therefore, serovar reassignment is considered a large source of uncertainty around the final data, and it is proposed that future models use a stochastic approach for reassigning, allowing this uncertainty to be expressed and quantified.

The consumption dataset presented a special challenge, as it had to be based on estimates from surrogate trade data, and an evaluation of the quality of EUROSTAT data revealed major inconsistencies in the intra-EU trade statistics [30]. However, according to comparison with WHO GEMS/Food, this approach produced valid results, as 19/24 countries had a PSI of $\geqslant 0{\cdot}9$ and three had a value of $>0{\cdot}8$, suggesting that the consumption profiles composed using EUROSTAT data are highly similar to the GEMS/Food profiles for most countries. An exception was noted for Cyprus, which is likely to be a reflection of the large proportion of extrapolated data, and which may have an effect on the attribution outcomes for that country. Nonetheless, the dataset as a whole showed 91% similarity.

In conclusion, as long as a thorough data evaluation is performed and specific countries and reservoirs with insufficiently representative data are excluded, public surveillance and monitoring data from multiple countries can potentially be used for scientific purposes, particularly for microbial subtyping-based source attribution methods. This could be a first step for the conduction of source attribution studies in countries or regions where no country-wide baseline studies have been conducted, but where programmes for *Salmonella* monitoring in food or surveillance in humans are currently up and running.

ACKNOWLEDGEMENTS

We acknowledge Timour Koupeev from Vose Risk Consulting for the collaboration on the management of the EUROSTAT data.

The staff of EFSA's Task Force of Zoonoses Data Collection is acknowledged for providing the original datasets necessary to conduct this study. The views or positions expressed in this publication do not necessarily represent in legal terms the official position of the European Food Safety Authority. The European Food Safety Authority assumes no responsibility or liability for any errors or inaccuracies that may appear.

The views and opinions of the authors expressed herein do not necessarily state or reflect those of the ECDC. The accuracy of the authors' statistical analysis and the findings they report are not the responsibility of ECDC. ECDC is not responsible for conclusions or opinions drawn from the data provided. ECDC is not responsible for the correctness of the data and for data management, data merging and data collation after provision of the data.

ECDC shall not be held liable for improper or incorrect use of the data.

DECLARATION OF INTEREST

The initial version of the EU model was developed with partial funding from contract CT/EFSA/Zoonoses/2010/02 (contract value 45000 Euros) between EFSA and the DTU National Food Institute, in relation to Question no. EFSA-Q-2010-00685.

REFERENCES

1. **Greger M.** The human/animal interface: emergence and resurgence of zoonotic infectious diseases. *Critical Reviews in Microbiology* 2007; **33**: 243–299.
2. **World Health Organization.** The world health report 2007 (http://www.who.int/whr/2007/en/index.html). Geneva, Switzerland: World Health Organization. Accessed 7 July 2011.
3. **European Food Safety Authority.** The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2009. *EFSA Journal* 2011; **9**: 2090.
4. **Kuchenmüller T, *et al.*** Estimating the global burden of foodborne diseases – a collaborative effort. *Eurosurveillance* 2009; **14**: 1–4.
5. **Pires SM, *et al.*** Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathogens and Disease* 2009; **6**: 417–424.
6. **European Food Safety Authority.** EFSA Panel on biological hazards (BIOHAZ). Scientific opinion on a request from EFSA on overview of methods for source attribution for human illness from food borne microbiological hazards. *EFSA Journal* 2008; **764**: 1–43.
7. **Hald T, *et al.*** A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Analysis* 2004; **24**: 255–269.
8. **European Food Safety Authority.** Report 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.
9. **European Food Safety Authority.** Report of the task force on zoonoses data collection on the analysis of the baseline survey on the prevalence of *Salmonella* in slaughter pigs in the EU, 2006–2007. Part A: *Salmonella* prevalence estimates. *EFSA Journal* 2008; **135**: 1–111.
10. **European Food Safety Authority.** Report of the task force on zoonoses data collection on the analysis of the baseline survey on the prevalence of *Salmonella* in turkey flocks, in the EU, 2006–2007. Part A: *Salmonella* prevalence estimates. *EFSA Journal* 2008; **134**: 1–224.
11. **European Food Safety Authority.** Report of the task force on zoonoses data collection on the analysis of the baseline survey on the prevalence of *Salmonella* in broiler flocks of *Gallus gallus* in the EU, 2005–2006. Part A: *Salmonella* prevalence estimates. *EFSA Journal* 2007; **98**: 1–85.
12. **European Food Safety Authority.** EFSA Panel of Biological Hazards (BIOHAZ). Scientific Opinion on a request from European Commission on quantitative estimation of the impact of setting a new target for the reduction of *Salmonella* in breeding hens of *Gallus gallus*. *EFSA Journal* 2009; **1036**: 1–68.
13. **European Food Safety Authority.** EFSA Panel of Biological Hazards (BIOHAZ). Scientific Opinion on a quantitative estimation of the public health impact of setting a new target for the reduction of *Salmonella* in broilers. *EFSA Journal* 2011; **9**: 2106.
14. **European Food Safety Authority.** EFSA Panel of Biological Hazards (BIOHAZ). Scientific Opinion on an estimation of the public health impact of setting a new target for the reduction of *Salmonella* in turkeys. *EFSA Journal* 2012; **10**: 2616.
15. **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 Diseases* 2010; **7**: 143–151.
16. **Wahlström H, *et al.*** Source attribution of human *Salmonella* cases in Sweden. *Epidemiology and Infection* 2011; **139**: 1246–1253.
17. **Pires SM,** *Attributing human salmonellosis and campylobacteriosis to food, animal and environmental sources* (thesis). Copenhagen, University of Copenhagen, 2009, 157 pp.
18. **de Knegt LV,** *A multi-country approach for attributing human salmonellosis to animal reservoirs: global perspectives and application of surveillance data from the European Union* (thesis). Kongens Lyngby, Denmark: Technical University of Denmark, 2013, 109 pp.
19. **European Centre for Disease Prevention and Control.** Surveillance of communicable diseases in the European Union: a long-term strategy 2008–2013 (http://ecdc.europa.eu/en/aboutus/Key%20Documents/08–13_KD_Surveillance_of_CD.pdf). Solna, Sweden: European Centre for Disease Prevention and Control, 2007. Acessed 2 November 2012.
20. **European Food Safety Authority.** Analysis of the baseline survey on the prevalence of *Campylobacter* in broiler batches and of *Campylobacter* and *Salmonella* on broiler carcasses in the EU, 2008 – Part A: *Campylobacter* and *Salmonella* prevalence estimates. *EFSA Journal* 2010; **8**: 1503.
21. **European Food Safety Authority.** Analysis of the baseline survey on the prevalence of *Salmonella* in holdings with breeding pigs in the EU, 2008 – Part A: *Salmonella* prevalence estimates. *EFSA Journal* 2009; **7**: 1377.
22. **Statistical Office of the European Union (EUROSTAT).** (epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database). Accessed 25 March 2011.
23. **Wheeler JG, *et al.*** Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *British Medical Journal* 1999; **318**: 1046–1050.

24. **Havelaar AH, *et al.*** Estimating the true incidence of campylobacteriosis and salmonellosis in the European Union, 2013. *Epidemiology and Infection* 2012; **141**: 293–302.
25. **World Health Organization Collaborating Center for Reference and Research on *Salmonella*.** *Antigenic Formulae of the* Salmonella *Serovars*, 9th edn. Paris, France: Institut Pasteur, 2007.
26. **World Health Organization Global Foodborne Infections Network Country Databank (WHO/GFN CDB).** (http://thor.dfvf.dk/gss). Accessed 15 November 2010.
27. **European Food Safety Authority.** The community summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2007. *EFSA Journal* 2009; **223**: 312.
28. **European Food Safety Authority.** The community summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2008. *EFSA Journal* 2010; **8**: 1496.
29. **Association of Poultry Processors and Poultry Trade in the EU Countries.** Annual Report 2009. Brussels, Belgium: Association of Poultry Processors and Poultry Trade in the EU Countries.
30. **European Food Safety Authority.** Collection and routine analysis of import surveillance data with a view to identification of emerging risks. *EFSA Journal* 2010; **8**: 1531.
31. **World Health Organization Global Environment Monitoring System Food Consumption Cluster Diets.** GEMS/Food regional diets: regional per capita consumption of raw and semi-processed agricultural commodities, prepared by the Global Environment Monitoring System/Food Contamination Monitoring and Assessment Programme (GEMS/Food). Geneva, Switzerland: World Health Organization 2006.
32. **Rosef O, *et al.*** Serotyping of *Campylobacter jejuni, Campylobacter coli* and *Campylobacter laridis* from domestic and wild animals. *Applied Environmental Microbiology* 1985; **49**: 1507–1510.
33. **Feinsinger P, Spears EE, Poole RW.** A simple measure of niche breadth. *Ecology* 1981; **62**: 27–32.
34. **Mullner P, *et al.*** Source attribution of food-borne zoonoses in New Zealand: a modified Hald model. *Risk Analysis* 2009; **29**: 970–984.
35. **Mullner P, *et al.*** Molecular epidemiology of *Campylobacter jejuni* in a geographically isolated country with a uniquely structured poultry industry. *Applied Environmental Microbiology* 2010; **76**: 2145–2154.
36. **European Food Safety Authority.** EFSA Panel on biological hazards (BIOHAZ). Scientific Opinion on monitoring and assessment of the public health risk of '*Salmonella* Typhimurium-like' strains. *EFSA Journal* 2010; **8**: 1826.
37. **European Food Safety Authority.** The community summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2006. *EFSA Journal* 2007; **130**: 352.
38. **David J,** Attribution of human cases of salmonellosis to different species of production animals in France: adaptability and robustness of the microbial subtyping source attribution Bayesian model (thesis). Vannes, France: Université Européenne de Bretagne, 2009, 217 pp.