



HAL
open science

**Attribution des cas de salmonelloses humaines aux
différentes filières de production animale en France.
Adaptabilité et robustesse du modèle bayésien
d'attribution par typage microbiologique**

Julie David

► **To cite this version:**

Julie David. Attribution des cas de salmonelloses humaines aux différentes filières de production animale en France. Adaptabilité et robustesse du modèle bayésien d'attribution par typage microbiologique. Biologie cellulaire. Agrocampus - Ecole nationale supérieure d'agronomie de rennes, 2009. Français. NNT: . tel-00485441

HAL Id: tel-00485441

<https://theses.hal.science/tel-00485441>

Submitted on 20 May 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° Ordre : 2009-25
N° Série : B-201

THESE / AGROCAMPUS OUEST

Sous le sceau de l'Université Européenne de Bretagne
pour obtenir le diplôme de :

**DOCTEUR DE L'INSTITUT SUPERIEUR DES SCIENCES AGRONOMIQUES,
AGRO-ALIMENTAIRES, HORTICOLES ET DU PAYSAGE**

Spécialité : Biologie et Agronomie

Ecole Doctorale : Vie Agro Santé

présentée par :

Julie DAVID

**Attribution des cas de salmonelloses humaines aux différentes filières de
production animale en France**

Adaptabilité et robustesse du modèle bayésien d'attribution par typage
microbiologique

soutenue le 14 décembre 2009 devant la commission d'Examen

Composition du jury :

Mme Catherine DISENHAUS
M. Jean-Claude DESENCLOS
M. André RAVEL
M. Didier GUILLEMOT
M. Pascal SANDERS
Mme Claire DANE
Mme Caroline LELAIDIER

Président
Rapporteur
Rapporteur
Directeur de thèse
Directeur de thèse
Membre
Membre

Remerciements

Je tiens ici à remercier vivement

Mes directeurs de thèse, Didier Guillemot et Pascal Sanders pour m'avoir proposé ce sujet ardu et passionnant et m'avoir fait confiance tout au long de ces quatre années ;

M. André Ravel, professeur et épidémiologiste à l'Agence de Santé Publique du Canada et M. Jean-Claude Désenclos, Directeur scientifique de l'Institut de Veille Sanitaire pour m'avoir fait l'honneur d'accepter la charge de rapporteurs de cette thèse et avoir participé au jury malgré la distance, ainsi que pour leur lecture critique constructive du manuscrit ;

Catherine Disenhaus, professeur à l'Agrocampus Ouest pour m'avoir fait l'honneur d'accepter de présider le jury de thèse, Claire Dane, Vétérinaire en Chef à la DGA et Caroline Lelaidier, Inspecteur de la Santé Vétérinaire, enseignante à l'Agrocampus Ouest pour m'avoir fait l'honneur de juger mes travaux et de participer au jury de thèse ;

La Délégation Générale de l'Armement pour avoir financé en grande partie ces travaux ;

Tous les partenaires du projet, Centre National de Référence des salmonelles, unité EAZ du département maladies infectieuses de l'Institut de Veille Sanitaire, ainsi que toutes les unités de l'Afssa concernées (Lerqap/CEB, Derns/Paser à Maisons Alfort, HQPAP, EBEAC et EBEP à Ploufragan, BBSV à Lyon) pour leur implication, leurs conseils précieux et leur soutien ;

Les membres du groupe de travail européen sur l'attribution des salmonelloses et campylobactérioses du réseau d'excellence MedVetNet pour les réflexions communes et les discussions très instructives sur ce sujet ;

Laurence Watier (Institut Pasteur / Inserm), Claire Chauvin (Afssa) et Corinne Danan (Afssa) qui m'ont fait bénéficier de leur savoir-faire, m'ont guidée et soutenue tout au long de cette thèse; merci pour tout ce que vous m'avez apporté professionnellement et humainement et pour votre disponibilité jusqu'à la dernière minute ;

Mes collègues et amis fougerais et pasteuris, parce qu'il n'y a pas que le travail dans la vie, merci pour votre amitié, votre soutien, le coach surfing lors de mes passages parisiens, les discussions scientifiques (tout de même !!!) et tous les bons moments passés... et à venir ?!

Mathilde et Arzel, Jérôme, Valérie, Karine, Perrine, Alexia, Catherine, Sylvie and co, Isabelle et Hubert, ma famille fougeraise, pilier quotidien et attentionné, vous me manquez déjà ;

Jean-Claude, merci pour tout ;

Christine, merci pour tes conseils et ton amitié qui ont une très grande importance pour moi ;

Ma famille enfin (Olivier inclus) et mes amis, rochers dans la tempête ; je ne serais pas là sans vous.

Table des abréviations

95% CI	95% Credibility Intervals
ACM	Analyse en Correspondances Multiples
ADN	Acide Désoxyribonucléique
AFSSA	Agence Française de Sécurité Sanitaire des Aliments
AMR	Antimicrobial Resistance
Amx	Amoxicilline
ARN	Acide Ribonucléique
BLSE	Beta-lactamases à spectre étendu
C	Chloramphénicol
C3G	Céphalosporines de Troisième Génération
CA-SFM	Comité de l'Antibiogramme de la Société Française de Microbiologie
Caz	Ceftazidime
CDC	Center for Disease Control and Prevention
CMI	Concentration Maximale Inhibitrice
CNR	Centre National de Référence des Salmonelles
CREDOC	Centre de Recherche pour l'Etude et l'Observation des Conditions de vie
DDASS	Direction Départementale des Affaires Sanitaires et Sociales
DDSV	Direction départementale des Services Vétérinaires
DGAL	Direction Générale de l'Alimentation
DGCCRF	Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes
ECDC	European Centre for Disease Prevention and Control
EFSA	European Food Safety Authority
EU	European Union
Eucast	European Committee on Antimicrobial Susceptibility Testing
FBO	Food Business Operators
Gm	Gentamicine
HACCP	Hazard Analysis Critical Control Points
ICMSF	International Commission on Microbiological Specifications for Foods
INSERM	Institut National de la Santé et de la Recherche Médicale
INCA	Enquêtes Individuelles Nationales de Consommations Alimentaires
INSEE	Institut national de la statistique et des études économiques
InVS	Institut de Veille Sanitaire
K	Kanamycine
LERQAP	Laboratoire d'Etudes et de Recherches sur la Qualité des Aliments et les Procédés Agroalimentaires
LVD	Laboratoires Vétérinaires Départementaux
MCA	Multiple Correspondence Analysis (voir ACM)
MCSI	Mission de Coordination Sanitaire Internationale
MLST	Multilocus Sequence Typing
MLVA	Multiple Loci VNTR Analysis" ou "Analyse de plusieurs locus VNTR »
Na	Acide Nalidixique
NAC	Nouveaux Animaux de Compagnie
NRC	National Reference Centre (voir CNR)
NRL	National Reference Laboratory
OFIMER	Office national interprofessionnel des produits de la mer et de l'aquaculture

OIE	Office International des Epizooties
PCR	Polymerase Chain Reaction
PFGE	Pulse Field Gel Electrophoresis
RAPD	Random Amplification Polymorphic DNA
RASFF	Rapid Alert System for Food and Feed
RESSAB	Réseau d'Epidémio-Surveillance des Salmonelloses Bovines
RFLP	Restriction Fragment Length Polymorphism
RNOEA	Réseau National d'Observations Epidémiologiques en Aviculture
S	Streptomycine
SE	<i>Salmonella</i> sérovar Enteritidis
SN	<i>Salmonella</i> Network
SNP	Single Nucleotide Polymorphisms
SSS	Sulfamides
ST	<i>Salmonella</i> sérovar Typhimurium
Sxt	Sulfamides-triméthoprim
Te	Tétracycline
tec	Tonnes équivalent carcasse ou tonnes équivalent coquilles
TIA (C)	Toxi-Infections Alimentaires (Collectives)
UK	United Kingdom
VIH	Virus de l'Immunodéficience Humaine
VNTR	Variable Number of Tandem Repeat
WHO	World Health Organization

Tables des matières

Table des abréviations.....	1
Table des illustrations.....	6
Introduction.....	11
Objet et contexte de l'étude.....	13
1 Salmonelles et salmonellose	15
1.1 Salmonella	15
1.1.1 Caractéristiques bactériologiques.....	15
1.1.2 Nomenclature	17
1.2 Epidémiologie	18
1.2.1 Réservoirs	18
1.2.2 Voies de transmission	20
1.2.3 Dynamique de la contamination humaine	21
1.3 Salmonellose non typhique	22
1.3.1 Pathologie	22
1.3.2 Traitement	23
2 Les toxi-infections alimentaires à salmonelles, un enjeu majeur	24
2.1 Enjeux sanitaires	24
2.1.1 Les toxi-infections alimentaires, un problème persistant	24
2.1.2 Un pathogène d'importance	25
2.1.3 Evolution des infections à salmonelles.....	28
2.2 Enjeux socio-économiques.....	29
2.3 Enjeux commerciaux	30
3 La lutte contre les salmonelles	31
3.1 Elevage.....	32
3.2 Abattage et transformation	32
3.3 Distribution	33
3.4 Consommation	33
3.5 Stratégie de lutte et outils de gestion	33
4 Méthode d'attribution des salmonelloses humaines aux sources animales	34
Premier article	35
5 L'attribution par typage microbiologique	37
5.1 Pourquoi choisir cette méthode	37
5.2 Principe de la démarche et utilisation du bayésien et de la modélisation bayésienne	
5.3 Le sous-typage	40
5.3.1 Méthodes phénotypiques	40
5.3.2 Méthodes génotypiques	41
5.3.3 Utilisation en attribution.....	43
Problématique.....	47
Travaux personnels.....	49
Première partie : les données.....	51
Chapitre I : Le système de surveillance français des salmonelles (deuxième article).....	55
1 Architecture of the <i>Salmonella</i> surveillance system	61
1.1 Human sector	61
1.1.1 European regulatory framework	61
1.1.2 National surveillance system (table 1)	61
1.2 Agro-food sector	62
1.2.1 European regulatory framework	62

1.2.2 National surveillance system	64
2 Discussion/ conclusions	68
Chapitre II : Sélection des données et adaptation à l'utilisation en attribution.....	83
1 Données humaines	86
1.1 Données disponibles	86
1.2 Adaptation des données pour l'attribution	87
1.3 Résultats.....	88
2 Données animales	90
2.1 Données disponibles	90
2.1.1 Données de surveillance active.....	90
2.1.2 Données de surveillance passive.....	91
2.2 Adaptation des données pour l'attribution	93
2.2.1 Données de surveillance active.....	93
2.2.2 Données de surveillance passive.....	93
2.3 Résultats.....	94
2.3.1 Données de surveillance active.....	94
2.3.2 Données de surveillance passive.....	95
3 Sous typage.....	98
3.1 Sous-typage par PFGE.....	98
3.1.1 Méthode.....	98
3.1.2 Résultats	99
3.2 Sous-typage par antibiotype.....	99
3.2.1 Méthode.....	99
3.2.2 Résultats pour Enteritidis	100
3.2.3 Résultats pour Typhimurium	102
4 Données de consommation	105
4.1 Données disponibles	105
4.2 Construction de l'indicateur.....	106
4.2.1 Indice en base 100.....	107
4.2.2 Obtention des données de consommation à partir d'INCA	107
5 Intégration des données pour la démarche d'attribution.....	110
5.1 Sélection des sérotypes à inclure pour la modélisation	110
5.2 Types spécifiques	111
6 Données sur les denrées importées.....	112
6.1 Importations et consommation française.....	112
6.2 Sources de données explorées	113
6.2.1 Surveillance des denrées importées en France	113
6.2.2 Utilisation des données de surveillance européennes.....	114
6.3 Conclusion	114
Deuxième partie : Le modèle d'attribution par typage microbiologique.....	117
Chapitre III : Application de l'attribution dans divers pays européens (troisième article)	121
.....	121
1 Introduction.....	127
2 Material and Methods.....	128
2.1 Model overview	128
2.2 Data used in the model	128
3 Results	132
4 Discussion.....	133
Chapitre IV : Sensibilité du modèle aux priors informatifs (quatrième article).....	141
1 Introduction.....	146
2 Material and methods	147
2.1 Data	147
2.2 Description of the models	149
2.2.1 Simple approach	149

2.2.2 Bayesian approach	150
2.3 Software used	154
3 Results	154
3.1 Data	154
3.2 Attribution.....	155
3.2.1 Adequacy of the model.....	156
3.2.2 Number of cases per source.....	157
3.2.3 Type and source dependant parameters.....	157
4 Discussion.....	159
Chapitre V : Impact de la qualité des données sur les résultats obtenus avec le modèle (cinquième article).....	171
1 Introduction.....	176
2 Materials and methods.....	178
2.1 Data	178
2.1.1 Human cases.....	178
2.1.2 Food-animal sources	178
2.1.3 Subtyping.....	180
2.1.4 Consumption	180
2.2 Models.....	181
2.2.1 Simple model.....	181
2.2.2 Bayesian model.....	182
3 Results	184
3.1 Data	184
3.1.1 Human cases.....	184
3.1.2 Datasets.....	184
3.1.3 Consumption data.....	187
3.2 Impact of the data characteristics on the attribution results using the simple model	188
3.2.1 Contamination of the sources: proportion vs prevalence	188
3.2.2 Increasing sources' number	188
3.3 Attribution results using Bayesian model	190
4 Discussion.....	193
Discussion et conclusion générales.....	209
Références.....	217
Annexes.....	229
Annexe 1.....	231
Annexe 2.....	237
Annexe 3.....	245
Annexe 4.....	249
Annexe 5.....	255
Annexe 6.....	259

Table des illustrations

Table des figures

Objet et contexte de l'étude

Figure 0.1 : *Salmonella* Typhimurium envahissant des cellules humaines, microscope électronique à balayage

Figure 0.2 : Coupe de salmonelle

Figure 0.3 : Détection de *Salmonella* spp. dans les aliments selon la méthode normalisée ISO 6579 :2002 – Mode opératoire

Figure 0.4 : Nombre de notifications auprès du RASFF (Rapid Alert System for Food and Feed) selon le type de produits pour *Salmonella* en 2008

Figure 0.5 : Pyramide du « burden of illness » utilisée aux Etats-Unis (FoodNet) pour évaluer le nombre de cas lié à des infections alimentaires

Figure 0.6 : Points d'intervention potentiels pour la lutte contre les salmonelles

Figure 0.7 : Arbre de décision pour le choix de la méthode utilisée

Deuxième article

Figure 1 : Organization of the *Salmonella* surveillance system in France in 2009

Chapitre II

Figure 2.1 : Résumé des critères d'exclusion et effectifs

Figure 2.2 : Répartition par sérotype des cas humains domestiques sporadiques non typhiques (n=9 076, données CNR de 2005)

Figure 2.3 : Répartition des souches par sérotype en fonction de la filière pour la surveillance active

Figure 2.4 : Répartition des souches par sérotype en fonction de la filière pour la surveillance passive

Figure 2.5 : Répartition des souches de sérotype Enteritidis par profil et par source de données

Figure 2.6 : Répartition des souches de sérotype Typhimurium par profil et par source de données

Figure 2.7 : Disponibilité sur le marché des produits animaux sous forme d'un indice en base 100, année de référence 1999, données OFIVAL

Troisième article

Figure 1: Estimated proportion of cases attributed to specific sources in the UK, Sweden, The Netherlands and Denmark

Quatrième article

Figure 1: Surveillance results, types repartition in the sources and among the human cases, p indicated prevalence of the sources, n indicated the total number of human cases.

Figure 2: Comparison of posterior predicted cases per type and observed ones

Figure 3: Attribution results for the simple model, the four parameterizations and Mullner's model

Figure 4: Posterior means and 95% CI of bacteria dependant parameters

Cinquième article

Figure 1: type distribution in the human cases and the sources according to active and passive surveillance

Figure 2: Impact of the indicator on the attribution results, simple model, active dataset

Figure 3: Impact of the number of sources included in the model on the attribution results, simple model, passive dataset

Figure 4: Plot of the observed and expected cases for the individual types for the active and the passive surveillance

Figure 5: Posterior means and 95% CI of attributed cases per source according to surveillance mode

Table des tables

Objet et contexte de l'étude

Table 0.1 : Caractères généraux et classification des *Salmonella*

Table 0.2 : Formules antigéniques des sérovars les plus fréquemment isolés de *Salmonella enterica* subsp *enterica*, extrait du tableau de Kauffmann-White

Table 0.3 : Facteurs de risque concernant les infections à salmonelles non typhiques

Table 0.4 : Symptomatologie des salmonelloses non typhiques

Table 0.5 : Nombre de foyers de TIAC, de cas, d'hospitalisations et de décès, selon l'agent étiologique confirmé ou suspecté entre 1996 et 2005

Table 0.6 : Survie après une infection à *Salmonella* Typhimurium selon le profil d'antibiorésistance, Danemark, 1995-1999

Table 0.7 : Comparaison des performances de systèmes de typage épidémiologique

Deuxième article

Table 1 : Surveillance of human cases in 2008

Table 2 : Surveillance of food-animals in 2008

Table 3 : Surveillance of food in 2008

Table 4 : Non-human passive surveillance in 2008

Chapitre II

Table 2.1 : Données de surveillance active

Table 2.2 : Données de surveillance passive

Table 2.6 : Caractéristiques des sous-types pour Enteritidis

Table 2.7 : Caractéristiques des sous-types déterminés par ACM pour Typhimurium

Table 2.8 : données de consommation

Table 2.9 : Données de consommation chez les adultes de 15 ans et plus

Table 2.10 : Données de consommation chez les enfants de 3 à 14 ans

Table 2.11 : Données de consommation utilisées dans le modèle

Table 2.12 : Règles de décision pour la sélection des sérotypes

Table 2.13 : Types spécifiques associés aux données de surveillance active

Table 2.14 : Types spécifiques associés aux données de surveillance passive

Table 2.15 : Données Office de l'élevage, 2005

Table 2.16 : Origine des importations

Troisième article

Table 1: Variables used in the different European models

Table 2: Sources included in the different European models

Table 3: Estimated proportion of cases of human salmonellosis attributed to specific sources in the UK, Sweden, the Netherlands and Denmark (95% CI)

Table 4: Estimated source- (a_j) and food-dependent factors (q_j) for the UK, Sweden, the Netherlands and Denmark, mean and 95% CI

Quatrième article

Table 1: Informative and non informative prior information for Hald's model

Table 2: Specific types characteristics

Table 3: Posterior percentages of expected cases per source for the Specific-Types 2 and Mullner models in comparison to the simple approach

Cinquième article

Table 1: Characteristics of the specific types

Table 2: Posterior means and 95% CI of the source-dependant parameters

Table 3: Top 10 of the estimated type dependant parameters for each dataset

Introduction

Dans les pays développés, la sécurité sanitaire des aliments est devenue un enjeu de santé publique majeur, comme en témoigne un certain nombre de scandales récents, encéphalite spongiforme bovine, dioxine, syndrome hémolytique et urémique, ainsi que les débats qui émergent depuis quelques années sur les OGM ou l'utilisation des antibiotiques en élevage. Le risque alimentaire est peu toléré par les populations et certains pathogènes alimentaires zoonotiques sont au centre de bien des préoccupations.

Parmi ceux-ci, *Salmonella* est, en 2007, la deuxième cause de toxi-infections alimentaires bactériennes en Europe. Une épidémie d'ampleur mondiale liée au sérovar Enteritidis, dont le réservoir privilégié reste la poule pondeuse, a commencé dans les années 80. Depuis, des clones multi-résistants ont émergé et leur impact sur la santé humaine en terme de morbidité et de mortalité a été documenté. A cet enjeu sanitaire s'ajoutent des enjeux économiques liés au coût sociétal des nombreux cas humains engendrés ainsi que des enjeux commerciaux liés à l'impact de la contamination par *Salmonella* sur les échanges de denrées animales.

En Europe, *Salmonella* est un exemple d'implication des pouvoirs publics dans la gestion d'un danger alimentaire. En effet, ce pathogène zoonotique alimentaire, inscrit dans la réglementation, fait l'objet à la fois d'efforts en terme de surveillance et de lutte intégrées, lutte à laquelle les professionnels de l'agro-alimentaire sont associés, depuis la production primaire (élevage) jusqu'à la transformation. Etant un pathogène ubiquitaire (il est présent dans la quasi-totalité des environnements et des filières animales), une lutte efficace passe par l'évaluation de la part relative des différentes sources dans la genèse des cas humains, afin d'allouer au mieux les efforts de maîtrise nécessaires. *Salmonella* est donc un sujet privilégié de la démarche d'attribution (ou « source attribution » selon l'expression anglo-saxonne), démarche qui vise à déterminer pour une infection alimentaire donnée, la part relative des cas humains liée à chaque source potentielle et qui permet d'identifier, hiérarchiser et évaluer de potentielles interventions de santé publique.

Ainsi, en France, grand pays producteur et exportateur de denrées animales, une politique de lutte volontaire a été mise en place en élevage dès les années 80 en amont des filières de production d'œufs de consommation puis de poulets de chair, réputées être des réservoirs

majeurs de cette zoonose. Cette mesure a permis de réduire le nombre de cas humains de façon significative. Néanmoins, les salmonelles restent en France la première cause bactérienne de toxi-infections alimentaires. L'utilisation de l'attribution permettrait d'identifier de potentielles interventions et de continuer à en évaluer l'efficacité à long terme. La possibilité de l'application d'une méthode d'attribution (attribution par typage microbiologique), déjà utilisée en Europe et qui focalise plus particulièrement sur le point d'intervention potentiel élevage, a donc fait l'objet de ces travaux de thèse.

Objet et contexte de l'étude

1 Salmonelles et salmonellose

1.1 *Salmonella*

En 1885, aux Etats-Unis, Theobald Smith, sous la direction du Dr Daniel Elmer Salmon, travaille sur l'efficacité d'un vaccin bactérien chez le porc et découvre ce qu'il pense être l'agent causal du choléra porcin. Il s'agit en fait d'une nouvelle espèce bactérienne, *Salmonella enterica* (Figures 0.1 et 0.2).



Figure 0.1 : *Salmonella* Typhimurium envahissant des cellules humaines, microscope électronique à balayage (source : U.S. National Institute of Health)

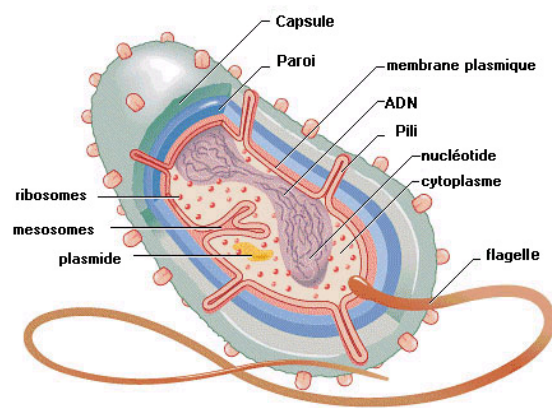


Figure 0.2 : Coupe de salmonelle (http://www.interet-general.info/article.php3?id_article=9632)

1.1.1 Caractéristiques bactériologiques

Salmonella est un bacille Gram négatif non sporulant, proche d'*Escherichia coli*, dont la mobilité propre est assurée par des flagelles péritriches (à l'exception de *S. Gallinarum* qui n'en possède pas) et qui est de type aéro-anaérobie. Ces bâtonnets de 2 à 3 μm de long sont des bactéries mésophiles, peu exigeantes d'un point de vue nutritionnel. Leur développement est optimal pour des températures proches de la température corporelle des animaux à sang chaud, 35 à 37°C, et un pH de 6,5 à 7,5. Leur multiplication reste assurée pour des températures de 6,7 à 41°C. Le large spectre de températures (-20 à 60°C) et de pH (4,1 à 9) auxquels elles sont capables de survivre, ainsi que leur capacité à résister à de a_w (activité de l'eau) de 0,94 en font des bactéries extrêmement résistantes aux conditions environnementales même difficiles (congélation) et expliquent leur caractère ubiquiste.

Les caractéristiques biochimiques spécifiques des salmonelles sont la réduction du nitrate en nitrite ; la possibilité de se contenter de citrate comme source de carbone ; pour le test des trois sucres : la fermentation du glucose mais pas du lactose, ni du saccharose ; la production de gaz à partir du glucose (à l'exception de *S. Typhi*) ; la synthèse de H₂S en milieu triple sucre et la réaction négative au test à l'oxydase. Les caractéristiques biochimiques des différentes espèces et sous-espèces¹ sont résumées dans la table 0.1 (Korsak 2004).

L'isolement et la caractérisation des salmonelles se fait en 4 étapes : pré-enrichissement (afin de revivifier les souches), enrichissement sur milieu sélectif (afin de favoriser la multiplication des salmonelles au détriment de la flore compétitrice), isolement en boîte de Pétri sur milieu sélectif et enfin identification sur gélose triple sucre. La méthode est adaptée en fonction de la matrice à examiner (aliment, échantillon environnemental, etc.) et fait l'objet de normes en constante évolution. La méthode normalisée ISO 6579:2002, de détection dans les aliments est présentée sur la figure 0.3.

Propriétés et caractères biochimiques	<i>Salmonella enterica</i>						<i>S. bongori</i>
	Habitat	Hommes et animaux à sang chaud		Animaux à sang froid et environnement			
Sous-espèce	<i>enterica</i>	<i>salamae</i>	<i>arizonae</i>	<i>diarizonae</i>	<i>houtenae</i>	<i>indiana</i>	
Nombre de sérovars (2005)	1504	502	95	333	72	13	22
<u>Caractères biochimiques :</u>							
ONPG (2h)	-	-	+	+	-	d	+
Gélatinase à 36°C	-	+	+	+	+	+	-
β-glucuronidase	d	d	-	+	-	d	-
γ-glutamyl transférase	d	+	-	+	+	+	+
Culture sur KCN	-	-	-	-	+	-	+
Dulcitol	+	+	-	-	-	d	+
Malonate	-	+	+	+	-	-	-
Galacturonate	-	+	-	+	+	+	+
L(+)-tartrate	+	-	-	-	-	-	-
Salicine	-	-	-	-	+	-	-
Sorbitol	+	+	+	+	+	+	-
Lyse par le phage O1	+	+	-	+	-	+	+

Table 0.1 : Caractères généraux et classification des *Salmonella* (Le Minor, Popoff et al. 1986)

+ : positif pour 90 à 100% des souches ; - : négatifs pour 90 à 100% des souches ;
d : variable selon les souches

¹ Cf. § 1.1.2. suivant relatif à la nomenclature.

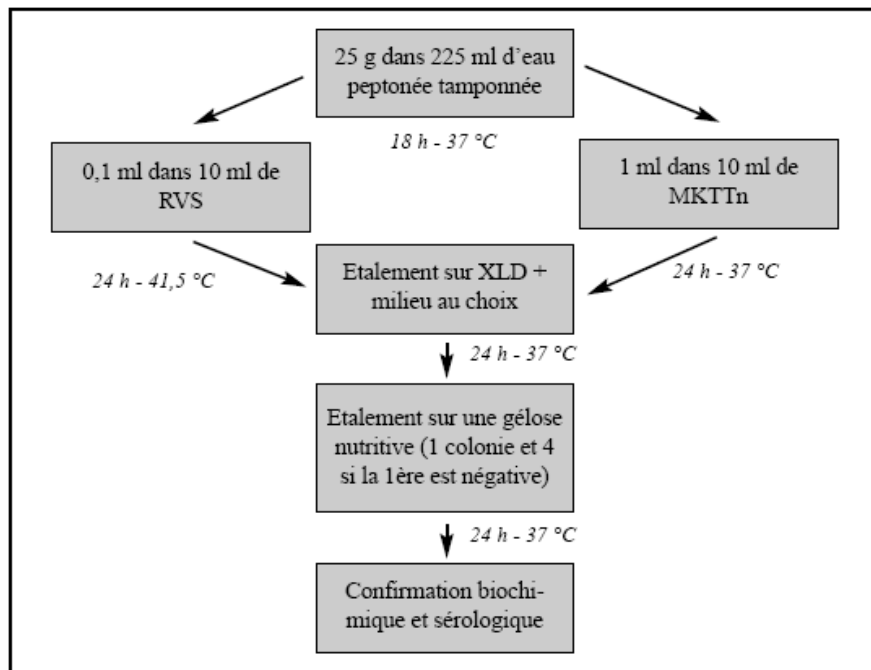


Figure 0.3 : Détection de *Salmonella* spp. dans les aliments selon la méthode normalisée ISO 6579 :2002 – Mode opératoire (Korsak 2004) (RVS, MKTTn, XLD désignent des milieux de culture de composition spécifique)

1.1.2 Nomenclature

Salmonella est une entérobactérie de la famille des *Enterobacteriaceae*. Le genre *Salmonella* comporte 3 espèces : *enterica*, *bongori* et *subterranea* (Le Minor and Popoff 1987), cette dernière espèce ayant été reconnue très récemment (Shelobolina, Sullivan et al. 2004). L'espèce *enterica*, qui possède un spectre d'hôtes très large, comprend elle-même 6 sous-espèces différenciées par leur biotype : *arizonae*, *diarizonae*, *enterica*, *houtenae*, *indica* et *salamae*. Les salmonelles sont enfin subdivisées selon leurs caractères antigéniques en plus de 2500 sérovars, ou sérotypes, dont la majeure partie appartient à la sous-espèce *S. enterica* subsp. *enterica* (table 0.1). Les sérovars sont définis selon les antigènes somatiques O (de nature polysaccharidique), flagellaires H (de nature protéique), et capsulaires Vi. Ces derniers sont rares, ils n'ont été identifiés que chez 3 sérovars, Typhi, paratyphi C et Dublin. Leurs formules simplifiées peuvent être trouvées dans le tableau de Kauffmann-White (Popoff and Le Minor 1997). Les sérovars relatifs à la sous-espèce *enterica*, portent un nom correspondant usuellement au lieu de leur premier isolement et s'écrivent avec une majuscule et en caractère romain (et non en italique). Ainsi, le sérotype Typhimurium par exemple, s'écrit *Salmonella enterica* subsp. *enterica* sérovar Typhimurium ou de façon plus concise *Salmonella* Typhimurium ou *S.* Typhimurium. Un extrait du tableau de Kauffmann-White est présenté pour les sérotypes de la sous-espèce *enterica* les plus fréquemment isolés (table 0.2).

L'adoption de cette nomenclature est le fruit d'un processus complexe, entre controverses et confusions, qui a donné lieu à un avis de la Commission Judiciaire recommandant d'utiliser le nouveau système présenté ici. Ce système est basé sur les nomenclatures validement publiées par l'Opinion judiciaire n°80, couplées à l'interprétation taxonomique de Le Minor et Popoff (1987) et à l'interprétation taxonomique de Reeves et al 1989 (Tindall, Grimont et al. 2005).

Groupe	Sérotype	Ag O	Ag H ₁	Ag H ₂
B (O : 4)	Typhimurium	<u>1</u> ,4,[5],12	i	1,2
	Saintpaul	<u>1</u> ,4,[5],12	e,h	1,2
	Brandenburg	<u>1</u> ,4,[5],12, <u>27</u>	l,v	e,n,z ₁₅
	Derby	<u>1</u> ,4,[5],12	f,g	[1,2]
C1 (O : 6,7)	Infantis	6,7, <u>14</u>	r	1,5
C2 (O : 6,8)	Bovismorbificans	6,8, <u>20</u>	r,[i]	1,5
	Hadar	6,8	z10	e,n,x
D (O : 9)	Enteritidis	<u>1</u> ,9,12	[f],g,m,[p]	-
	Dublin	<u>1</u> ,9,12[Vi]	g,p	-

Table 0.2 : Formules antigéniques des sérovars les plus fréquemment isolés de *Salmonella enterica* subsp *enterica*, extrait du tableau de Kauffmann-White (Le Minor and Popoff 1987; Popoff and Le Minor 1997), d'après (Brisabois 2001)

1.2 Epidémiologie

1.2.1 Réservoirs

Les salmonelles sont des pathogènes intestinaux (D'Aoust 1991). Présentes dans les intestins de l'homme et des animaux - leur réservoir principal - elles peuvent, suite à une contamination fécale, survivre dans l'environnement (eau et sol) plusieurs mois (Korsak 2004). Leur ubiquité se traduit par un large spectre de réservoirs : humains (Todd, Greig et al. 2008) et animaux, mammifères (Dechet, Scallan et al. 2006; Swanson, Snider et al. 2007), volatiles (Uyttendaele, Debevere et al. 1998; Hennessy, Cheng et al. 2004; Kimura, Reddy et al. 2004; Arsenault, Letellier et al. 2007b; Arsenault, Letellier et al. 2007c), reptiles (Woodward, Khakhria et al. 1997; De Jong, Andersson et al. 2005), crustacés (Butt, Aldridge et al. 2004)... Leur capacité de survie leur permet également de persister dans des réservoirs secondaires comme les boues d'épuration (Sahlstrom, de Jong et al. 2006), les aliments d'origine animale (Haeghebaert, Sulem et al. 2003; Oliver, Jayarao et al. 2005) ou végétale (Kirk, McKay et al. 2008), les fruits et légumes (Brandl 2006)... La variété de réservoirs potentiels est ainsi illustrée par la figure 0.4, qui présente le nombre de notifications de contamination par *Salmonella* recensées par le système européen d'alerte rapide en 2008

par type de source, et la table 0.3, qui synthétise les facteurs de risques d'infection à *Salmonella* les plus fréquemment évoqués dans la littérature selon le sérotype.

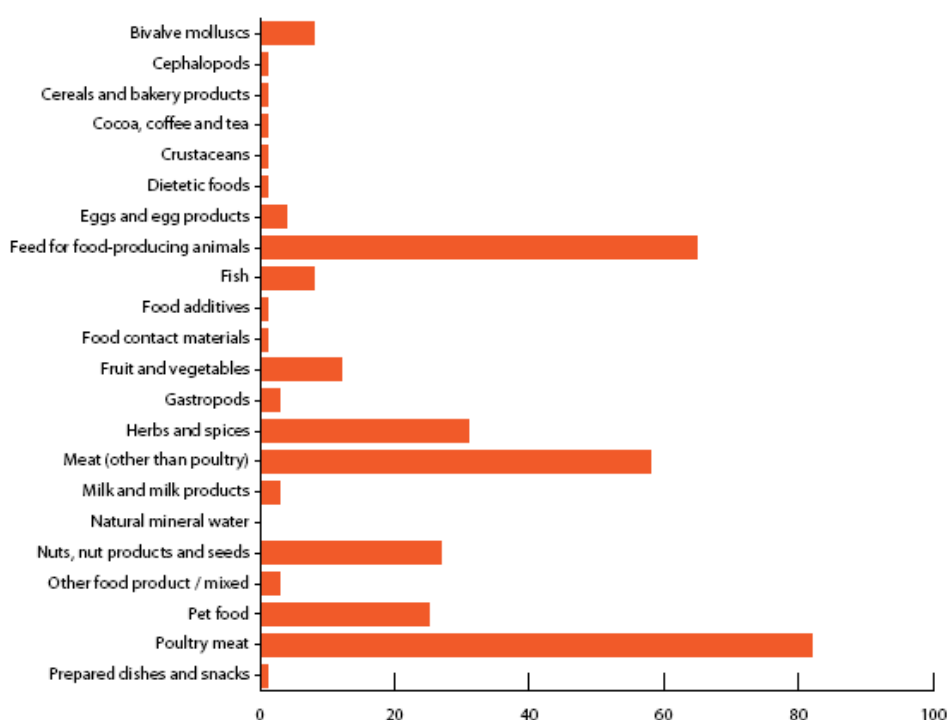


Figure 0.4 : Nombre de notifications auprès du RASFF (Rapid Alert System for Food and Feed) selon le type de produits pour *Salmonella* en 2008 (Anonymous 2009b)

<i>Salmonella enterica</i> serotype	Risk factors
<i>S. Typhimurium</i>	Consumption of undercooked meat or ground beef and dairy products, particularly eggs (raw eggs are associated with the highest risk); receipt of antibiotics, H2 antagonists, or proton pump inhibitors before exposure; residence on a farm with livestock; international travel; playing in a sand box
<i>S. Enteritidis</i>	Consumption of eggs, particularly outside of the home; international travel
<i>S. Newport</i>	Receipt of antibiotics before exposure; consumption of uncooked ground beef, runny scrambled eggs, or omelets prepared in the home; consumption of ground beef
<i>S. Heidelberg</i>	Consumption of poultry or eggs, especially consumption of eggs outside of the home
All <i>Salmonella</i> serotypes	In addition to the risk factors above, consumption of raw vegetables, including sprouts; nosocomial acquisition; exposure to reptiles (lizards, snakes, and turtles), amphibians (frogs and newts), baby poultry, and household pets fed natural pet treats, raw food diets, and other animal food; and cleaning aquaria in sinks
Antibiotic-resistant <i>Salmonella</i> strains ^a	Antibiotic use in animals, receipt of antibiotics before exposure

^a Most strains found in foods are antibiotic resistant.

Table 0.3 : Facteurs de risque concernant les infections à salmonelles non typhiques (DuPont 2007)

Comme le mentionne la table 0.3, certains sérotypes sont plus particulièrement présents chez certains hôtes, ils peuvent être ainsi classés selon l'espèce animale cible (Korsak 2004).

Un premier groupe de sérotypes particulièrement pathogènes pour l'homme (ils sont responsables de la fièvre typhoïde et paratyphoïde) n'est exclusivement isolé que chez celui-ci. On ne connaît pas de réservoir animal à *S. Typhi*, *S. Paratyphi* et *S. Sendai*. Ces sérotypes dits typhiques ne seront pas considérés dans ces travaux qui concernent les infections à salmonelles non typhiques dont l'origine est majoritairement alimentaire.

D'autres sérotypes sont spécifiques d'espèces animales, auxquelles ils sont étroitement adaptés et chez qui ils sont responsables de pathologies parfois graves (diarrhées, avortements...). On peut ici citer *Abortus ovis* chez les ovins, *Abortus equi* chez les équidés, *Gallinarum-pullorum* chez les volailles, *Dublin* chez les bovins et *Choleraesuis* et *Typhisuis* chez les porcins. Bien que fortement liés aux espèces citées, ces sérotypes sont transmissibles à l'homme pour qui ils peuvent être pathogènes, voire hautement pathogènes (D'Aoust 1991).

Enfin, la majorité des salmonelles sont ubiquistes. Elles franchissent la barrière d'espèce et peuvent se transmettre de l'animal à l'homme et réciproquement. Elles sont potentiellement pathogènes pour l'homme, à différents degrés (Jones, Ingram et al. 2008) et font la plupart du temps l'objet d'un portage sain chez l'animal voire d'une infection subclinique (Korsak 2004). Ainsi, dans la plupart des cas, l'infection à salmonelles d'un animal est silencieuse. La détection de la contamination nécessitera donc soit une recherche bactériologique sur prélèvement soit une recherche sérologique.

Ainsi, les salmonelles non typhiques sont des pathogène zoonotiques, c'est-à-dire transmissibles de l'animal à l'homme et vice-versa.

1.2.2 Voies de transmission

La principale voie de contamination pour l'homme est alimentaire (D'Aoust 1994; Angulo, Johnson et al. 2000). Mead (1999) estime en effet que l'alimentation est, aux Etats-Unis, la cause de 95% des infections à salmonelles (Mead, Slutsker et al. 1999). L'infection résulte alors de la consommation d'aliments contaminés. Pour ceux-ci, la contamination peut être intrinsèque, comme cela peut être le cas pour les œufs (Tauxe 1997; Rabsch, Tschäpe et al. 2001), ou secondaire, suite au contact avec des matières fécales lors de l'abattage pour les aliments issus d'animaux contaminés, ou encore avec une surface ou un autre aliment contaminé lors de la préparation ou de la transformation des denrées ; on parle alors de contamination croisée.

Cependant, la contamination peut également avoir lieu par contact avec des animaux infectés, notamment avec des animaux de compagnie et des NAC (nouveaux animaux de

compagnie) (Woodward, Khakhria et al. 1997; De Jong, Andersson et al. 2005; Swanson, Snider et al. 2007). Une équipe américaine a ainsi attiré l'attention sur la nécessité de tenir compte de ces voies de transmission non alimentaires dont l'importance serait sous-estimée selon eux (Barber, Miller et al. 2003). Une étude d'attribution récente (Evers, Van Der Fels-Klerx et al. 2008) concernant *Campylobacter*, autre pathogène zoonotique majeur, vient à l'appui de cette idée. Il y apparaît que la transmission par contact direct avec des animaux serait une voie majeure de transmission de ce pathogène. Ces derniers résultats sont cependant à prendre avec précaution en raison de fortes incertitudes liées à la méthode utilisée.

1.2.3 Dynamique de la contamination humaine

Les salmonelles sont à l'origine soit d'infections en apparence isolées, dites sporadiques, les plus nombreuses (Tauxe 1997), soit de phénomènes épidémiques : cas groupés ou foyers appelés encore toxi-infections alimentaires collectives (TIAC, définies comme au moins deux cas d'infection d'une même maladie survenant chez des personnes ayant partagé le même repas). Les TIAC ont évolué ces dernières décennies (Tauxe 1997), avec l'industrialisation et la globalisation des approvisionnements. Aux foyers restreints, liés à un aliment fortement contaminé consommé par un petit groupe de personnes (lors d'un rassemblement familial par exemple), se sont ajoutées des épidémies de plus grande ampleur, diffuses, pouvant impliquer plusieurs pays et des centaines de cas (Ammon and Tauxe 2007), et qui sont liées à une plus faible contamination d'aliments largement distribués. Ces dernières sont beaucoup plus difficiles à détecter.

Différents facteurs jouent sur la contamination humaine : dose ingérée (Bollaerts, Aerts et al. 2008), susceptibilité de la personne (degré d'immunité) (Flint, Van Duynhoven et al. 2005), mais également virulence et pathogénicité de la souche ingérée qui dépend notamment du sérotype (Coleman, Marks et al. 2004; Weinberger, Andorn et al. 2004; Foley and Lynne 2008; Jones, Ingram et al. 2008). Ainsi, la contamination d'un aliment seule ne suffit pas à prédire les infections qui vont résulter de sa consommation. Et ceci d'autant moins que tous les aliments ne vont pas avoir la même capacité à véhiculer les salmonelles ; celle-ci varie selon leurs propriétés physico-chimiques, selon le processus de fabrication mais aussi selon le mode de consommation (D'Aoust 1989). Ainsi, une mayonnaise à base d'œufs crus sera potentiellement plus « contaminante » qu'un morceau de bœuf consommé en ragoût. Globalement la relation dose-réponse (en terme d'infections), sera à considérer pour un couple sérotype-aliment (Bollaerts, Aerts et al. 2008).

1.3 Salmonellose non typhique

1.3.1 Pathologie

Les salmonelles non typhiques sont chez l'homme des pathogènes intracellulaires facultatifs, principalement d'origine alimentaire (Mead, Slutsker et al. 1999). Elles peuvent causer des gastroentérites aiguës avec céphalées, douleurs abdominales, nausées, vomissements, fièvre, dont l'évolution est le plus souvent favorable en quelques jours (InVS 2004) (Table 0.4). Cependant, chez les personnes fragiles, l'infection à salmonelles peut se traduire par une infection systémique sévère, usuellement de type intestinal, qui peut évoluer en infection extra-intestinale (ostéomyélite, bactériémie, infection urinaire...) notamment chez les patients immunodéprimés (Hohmann 2001). Les sérotypes Enteritidis et Typhimurium sont responsables de la majeure partie de ces infections particulièrement sévères (respectivement 26% et 50%). Dans le cas des infections extra-intestinales, le taux de décès est plus élevé, de l'ordre de 9% pour les enfants et de 17% pour les immunodéprimés (Ruiz, Rodriguez et al. 2004).

Pour les formes les plus courantes, l'excrétion après infection dure moins d'un an et dans la plupart des cas moins de trois mois. Cependant, il existe des porteurs chroniques (moins de 1% des infections), pour lesquels l'excrétion peut durer au-delà d'un an. Les salmonelles peuvent alors se nicher dans les schistosomes (parasites) ou dans des anomalies structurelles des systèmes urinaires et biliaires.

Symptômes	Salmonellose non typhique
Période d'incubation	8-72 heures
Diarrhée	+++ (liquide)
Douleurs abdominales	+++
Fièvre	+ (<48h)
Infection systémique	Rare (S. Choleraesuis, S. Dublin)
Durée	≤5 jours

Table 0.4 : Symptomatologie des salmonelloses non typhiques (D'Aoust 1991)

1.3.2 Traitement

Population générale

La plupart des infections à salmonelles ne nécessite pas de traitement antibiotique, elles relèvent essentiellement d'un traitement symptomatique, réhydratation orale principalement (Cezard, Chouraqui et al. 2002). Une méta-analyse de la collaboration Cochrane concernant le traitement des salmonelloses de l'adulte et de l'enfant bien portant montre que le traitement antibiotique des salmonelloses n'est associé à aucun bénéfice clinique et augmente les effets adverses et la durée d'excrétion (Sirinavin and Garner 2000). Les effets négatifs des traitements antibiotiques s'expliqueraient par la destruction de la flore de barrière intestinale. Cependant, le traitement antibiotique reste nécessaire chez les patients présentant un risque de développer des complications.

Population à risque

La population à risque de développer une bactériémie se compose des nouveau-nés de moins de trois mois voire moins d'un an, des personnes de plus de 50 ans et des patients dont le système immunitaire est affaibli (transplantés, immunodéprimés, maladies lymphoprolifératives, schistosomoses, HIV, joints prosthétiques, ...). Le traitement de première intention est alors un traitement oral ou en intraveineuse de 48 à 72 heures (ou jusqu'à tombée de la fièvre) d'ampicilline ou triméthoprime-sulfaméthoxazole si le germe est sensible. En cas de résistance à ces antibiotiques, les céphalosporines de troisième génération (C3G) ou les fluoroquinolones seront utilisées. Chez l'enfant, les fluoroquinolones étant contre-indiquées, elles ne seront prescrites qu'en cas de résistance aux C3G. Le traitement de première intention est court, afin d'éviter les risques de portage chronique et de récurrence associés aux traitements prolongés (Ruiz, Rodriguez et al. 2004; Stoycheva and Murdjeva 2006).

2 Les toxi-infections alimentaires à salmonelles, un enjeu majeur

2.1 Enjeux sanitaires

2.1.1 Les toxi-infections alimentaires, un problème persistant

L'importance des toxi-infections alimentaires (TIA) ne peut qu'être estimée, mais se mesure en millions de cas annuels, quel que soit le pays concerné (Flint, Van Duynhoven et al. 2005). Ces chiffres restent élevés malgré les mesures de prévention et de lutte mises en place, des classiques mesures d'hygiène (pasteurisation et réfrigération) aux stratégies de lutte intégrée (plans de contrôles européens en élevage), en passant par l'investigation d'épidémies et la mise en place de la méthode HACCP (Hazard Analysis Critical Control Points) chez les professionnels de l'agroalimentaire. En effet, un certain nombre d'éléments de contexte concourent à favoriser le succès de ces infections (Allard 2002; Desenclos, Vaillant et al. 2002; DuPont 2007; Sofos 2008). D'une part, Les modes de production industriels et la globalisation de l'approvisionnement conduisent en cas de contamination à une large diffusion des aliments contaminés et ainsi potentiellement à des foyers épidémiques de grande ampleur et diffus. L'évolution des habitudes alimentaires favorise également l'exposition aux pathogènes alimentaires, en s'orientant vers la consommation de produits crus ou peu cuits et la restauration hors-foyer qui est un facteur de risque connu, notamment pour les infections à salmonelles (Hennessy, Cheng et al. 2004; Jones and Angulo 2006). Il en va de même pour la circulation des aliments et des personnes, qui favorise l'exposition des populations à des pathogènes « exotiques ». D'autre part, le vieillissement de la population et l'efficacité des soins conduisent à l'augmentation de la population à risque, personnes âgées et fragilisées, telles que les malades du virus de l'immunodéficience humaine (VIH).

La persistance des TIA trouve donc probablement ses racines dans ces facteurs favorisant l'exposition d'une part et l'augmentation des populations vulnérables d'autre part. Enfin, l'impact de l'amélioration des systèmes de surveillance des cas de TIA sur les tendances mesurées ne doit pas être négligé.

2.1.2 Un pathogène d'importance

Dans les pays occidentaux, les salmonelles non typhiques comptent parmi les principaux pathogènes responsables des TIA évoquées ci-dessus. Elles sont en effet la deuxième cause bactérienne de gastro-entérites, après *Campylobacter*, en Europe (Anonymous 2009a) et aux Etats-Unis (Mead, Slutsker et al. 1999), où elles sont en revanche au premier rang en ce qui concerne les hospitalisations et les décès relatifs. Ainsi, 1,4 millions de cas pour 15 000 hospitalisations et 400 décès seraient dus chaque année à *Salmonella* sur le territoire américain (Voetsch, Van Gilder et al. 2004), alors que plus de 150 000 cas confirmés sont enregistrés en Europe (Anonymous 2009b). En Australie, *Salmonella* apparaît au deuxième rang en ce qui concerne les cas sporadiques et au premier rang des causes de foyers épidémiques (Anonymous 2005), pour un total estimé de 5,4 millions de cas chaque année (Flint, Van Duynhoven et al. 2005). Au Royaume-Uni, près de 107 000 cas seraient liés annuellement à *Salmonella*, qui y est la quatrième cause de gastro-entérites aiguës et se place au premier rang en terme de décès relatifs (Flint, Van Duynhoven et al. 2005).

En France enfin, l'Institut de Veille Sanitaire (InVS 2004) a évalué le nombre annuel de cas confirmés (sporadiques ou liés à des foyers) en 2004 entre 30 600 et 41 140 cas dont 92 à 535 décès, soit un taux de mortalité de 0,3 à 1,3%, ce qui en fait la première cause de gastro-entérites bactériennes d'origine alimentaire. Ce chiffre de quelques dizaines de milliers de cas peut paraître modeste, mais il ne s'agit que de la pointe émergée de l'iceberg. En effet, les cas confirmés correspondent à des malades ayant consulté un médecin pour la gastro-entérite, auxquels a été prescrite une recherche de salmonelles, dont les résultats (positifs) ont été enregistrés dans l'une des sources ayant servi à l'estimation faite par l'InVS. Si l'on se réfère à la pyramide décrivant la surveillance des infections d'origine alimentaire (Figure 0.5), la surveillance basée sur les laboratoires, tel que c'est le cas en France, ne touche que la pointe de la pyramide, menant à une probable forte sous-estimation du nombre réel de cas correspondants.

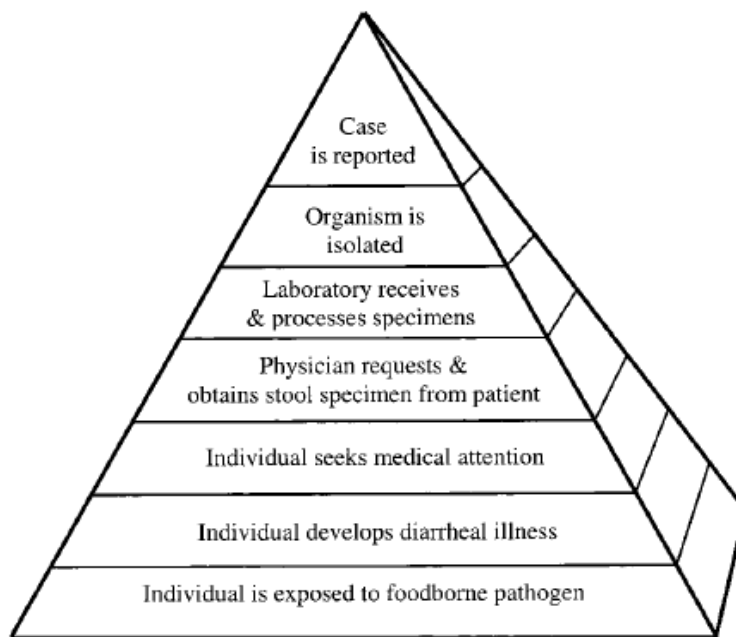


Figure 0.5 : Pyramide du « burden of illness » utilisée aux Etats-Unis (FoodNet) pour évaluer le nombre de cas lié à des infections alimentaires (Allos, Moore et al. 2004)

Si l'on s'intéresse plus particulièrement aux cas groupés (ou foyers), *Salmonella* se range en France au premier rang des agents confirmés lors de toxi-infections alimentaires collectives (Delmas, Gallay et al. 2007), que l'on considère le nombre de foyers, de cas, d'hospitalisation ou de décès pour la période 1996-2005 (Table 0.5).

Par ailleurs, les taux d'hospitalisation et de décès observés dans les différents pays et la place prédominante qu'y occupe *Salmonella* signent la sévérité des symptômes liés à cette bactérie, sévérité confirmée par ailleurs dans la littérature, que l'on considère les conséquences à court ou à long terme (Helms, Vastrup et al. 2003; Kennedy, Villar et al. 2004; Helms, Simonsen et al. 2006).

Agent	Foyers		Cas		Hospitalisations		Décès	
	N	% ^a	N	% ^a	N	% ^a	N	% ^b
Agents confirmés								
<i>Salmonella</i>	1 713	64,2 %	16 230	48,8 %	2 961	61,0 %	21	0,13 %
dont Enteritidis	936	54,6 %	9 152	56,4 %	1 759	10,8 %	13	0,14 %
Typhimurium	312	18,2 %	2 976	18,3 %	500	3,1 %	5	0,17 %
Autres sérotypes ^c	132	7,7 %	1 999	12,3 %	312	1,9 %	2	0,10 %
Sérotypes indéterminés	333	19,4 %	2 103	13,0 %	390	2,4 %	1	0,05 %
<i>Clostridium perfringens</i>	136	5,1 %	5 375	16,2 %	42	0,1 %	2	0,04 %
<i>Shigella</i>	42	1,6 %	337	1,0 %	58	0,2 %	0	-
<i>Campylobacter</i>	37	1,4 %	426	1,3 %	55		1	0,23 %
<i>Staphylococcus aureus</i>	366	13,7 %	5 750	17,3 %	1 182	3,6 %	2	0,03 %
<i>Bacillus cereus</i>	94	3,5 %	1 766	5,3 %	148	0,4 %	4	0,23 %
Histamine	89	3,3 %	777	2,3 %	149	0,4 %	0	-
Virus	38	1,4 %	950	2,9 %	3	0,0 %	0	-
Autres pathogènes ^d	152	5,7 %	1 622	4,9 %	258	0,8 %	2	0,12 %
Total agents confirmés	2 667	45,6 %	33 233	100,0 %	4 856	14,6 %	32	0,10 %
Agents suspectés								
<i>Salmonella</i>	261	12,6 %	3 558	11,4 %	316	1,0 %	1	0,03 %
<i>Clostridium perfringens</i>	383	18,5 %	8 956	28,8 %	65	0,2 %	3	0,03 %
<i>Shigella</i>	3	0,1 %	20	0,1 %	1	0,0 %	0	-
<i>Campylobacter</i>	10	0,5 %	250	0,8 %	15	0,0 %	0	-
<i>Staphylococcus aureus</i>	744	35,9 %	8 926	28,7 %	812	2,6 %	0	-
<i>Bacillus cereus</i>	196	9,5 %	3 532	11,4 %	225	0,7 %	0	-
Histamine	143	6,9 %	926	3,0 %	162	0,5 %	1	0,11 %
Virus	191	9,2 %	3 759	12,1 %	47	0,2 %	0	-
Autres pathogènes ^e	143	6,9 %	1 166	3,8 %	102	0,3 %	2	0,17 %
Total agents suspectés	2 074	35,5 %	31 093	38,7 %	1 745	5,6 %	7	0,02 %
Total agents indéterminés	1 106	18,9 %	16 025	19,9 %	763	2,5 %	6	0,04 %
Total foyers	5 847	100,0 %	80 351	100,0 %	7 364	100,0 %	45	100,0 %

^a Pour les différents agents, % par rapport au total des agents confirmés ou suspectés. Pour les sérotypes de *Salmonella*, % par rapport au total des *Salmonella*.

^b Nombre de décès pour 100 malades pour chaque germe.

^c *S. Hadar* (21 foyers) - *S. Heidelberg* (19 foyers) - *S. Virchow* (19 foyers) - *S. Newport* (10 foyers) – Autres sérotypes (63 foyers).

^d *E. Coli* (44 foyers) - *C. Botulinum* (22 foyers) - Coliformes (18 foyers) - *V. Parahaemolyticus* (12 foyers) - DSP (20 foyers) - Toxique (6 foyers) - VTEC O 157 (4 foyers) - VTEC O 148 (1 foyer) - *Streptococcus* (2 foyers) - *Trichinella* (4 foyers) - *Bruceella* (1 foyer) - Levures (1 foyer) - Toxopl. (1 foyer) - VHA (1 foyer) - *Yersinia enterocolitica* (1 foyer) - Autre (14 foyers).

^e *E. Coli* (46 foyers) - *Ciguatera* (39 foyers) - Agent allergisant (10 foyers) - DSP (20 foyers) - *C. botulinum* (5 foyers) - Toxique (5 foyers) - *V. Parahaemolyticus* (2 foyers) - Champignons (1 foyer) - Levures (1 foyer) - Autre (14 foyers).

Table 0.5 : Nombre de foyers de TIAC, de cas, d'hospitalisations et de décès, selon l'agent étiologique confirmé ou suspecté entre 1996 et 2005 (Delmas, Gallay et al. 2007)

2.1.3 Evolution des infections à salmonelles

Salmonella est un danger en constante évolution, doté d'une forte capacité adaptative. Dans les dernières années, deux phénomènes majeurs illustrent cela (Rabsch, Tschäpe et al. 2001; Velge, Cloeckart et al. 2005).

Ainsi, dans les années 80, émerge aux Etats-Unis et en Europe le sérotype Enteritidis qui est particulièrement associé aux espèces aviaires et a la capacité de provoquer une infection trans-ovarienne et par conséquent de contaminer les œufs. Rapidement, l'épidémie devient une pandémie mondiale et depuis des années, Enteritidis caracole en tête des classements de sérotypes parmi les cas humains (Rabsch, Tschäpe et al. 2001). C'est également le sérotype le plus souvent identifié lors de cas groupés ou foyers épidémiques (Delmas, Gallay et al. 2007; Greig and Ravel 2009). Les raisons de l'émergence et du succès fulgurant de ce sérotype ne sont à l'heure actuelle pas connues. L'hypothèse d'introduction d'une souche d'Enteritidis par un rodenticide (*Salmonella* étant utilisée comme agent biologique pour lutter contre les rongeurs) a été écartée et les hypothèses vont plutôt vers l'occupation d'une niche écologique restée vacante suite à l'éradication des salmonelles de sérotype Gallinarum chez le poulet (Rabsch, Tschäpe et al. 2001). Il reste que ce sérotype particulièrement pathogène demeure le plus représenté parmi les cas humains avec Typhimurium dans la majorité des pays occidentaux (Galanis, Lo Fo Wong et al. 2006; Maraki, Samonis et al. 2006; Kirk, McKay et al. 2008; Anonymous 2009a).

Plus récemment, *Salmonella* a acquis le statut de risque émergent par l'acquisition de résistances aux antibiotiques (Sofos 2008). Deux points sont ici en cause : d'une part la multi-résistance, et d'autre part l'émergence et la diffusion de résistances aux quinolones (Molbak, Gerner-Smidt et al. 2002b) et aux céphalosporines de troisième génération (C3G) (Miriagou, Tassios et al. 2004; Devasia, Varma et al. 2005), classes d'antibiotiques critiques pour les traitements thérapeutiques chez les populations à risques et notamment les enfants (DuPont 2007). En effet, outre le risque d'échec thérapeutique en cas de résistance, il a été montré que la multi-résistance et/ou les résistances aux quinolones et C3G sont associées chez les salmonelles à une morbidité et une mortalité accrues (Table 0.6) (Helms, Vastrup et al. 2002; Helms, Simonsen et al. 2004; Martin, Fyfe et al. 2004; Molbak 2005; Varma, Mølbak et al. 2005). Les mécanismes qui y sont associés seraient un avantage sélectif en cas de traitement antibiotique préalable (Glynn, Reddy et al. 2004), et une virulence accrue liée à des co-sélections de gènes de virulence avec les gènes de résistance (Barza 2002; Martinez and Baquero 2002; Molbak 2005; Foley and Lynne 2008). La multi-résistance, associée ou non aux résistances aux fluoroquinolones et C3G, est particulièrement associée au sérotype Typhimurium et au clone DT104, qui représente une part importante des souches de

Typhimurium (Rabatsky-Ehr, Whichard et al. 2004). Pour ce clone, les gènes de la penta-résistance sont situés sur le chromosome, dans l'îlot génomique SGI1 (Velge, Cloeckart et al. 2005), et non sur des éléments génétiques mobiles, et la multi-résistance y est donc très stable.

Resistance pattern	No. of patients	Relative mortality compared with the general population (95 % CI)
Pansusceptible	953	2.3 (1.5–3.5)
ACSSuT	283	4.8 (2.2–10.2)
Nx ^a	83	10.3 (2.8–37.8)
ACSSuTNx	40	13.1 (3.3–51.9)

NOTE. Adapted from [38]. All estimates are hazard ratios, calculated by conditional proportional hazard regression analysis, compared with an age- and sex matched sample of the general Danish population, and adjusted for underlying illness. A, resistant to ampicillin; C, resistant to chloramphenicol; Nx, resistant to nalidixic acid; S, resistant to streptomycin; Su, resistant to sulfonamides; T, resistant to tetracycline.

^a Resistant to at least nalidixic acid (i.e., first-generation quinolones); these strains have reduced susceptibility to fluoroquinolones, such as ciprofloxacin.

Table 0.6 : Survie après une infection à *Salmonella* Typhimurium selon le profil d'antibiorésistance, Danemark, 1995-1999 (Molbak 2005)

Ainsi, les salmonelles restent une cause majeure de gastro-entérites sévères, malgré un nombre de cas stables (Anonymous 2005) voire en décroissance (Anonymous 2009a) sur les dernières années. Elles représentent de plus un danger en constante mutation. L'incidence importante des salmonelloses ainsi que l'émergence de la multi-résistance et de résistances à des antibiotiques critiques en thérapeutique humaine, dans un contexte où le danger de l'antibiorésistance fait l'objet de toutes les attentions, notamment en France (Sabuncu, David et al. 2009), en font un enjeu sanitaire de premier ordre.

2.2 Enjeux socio-économiques

La salmonellose est une maladie à forte incidence, responsable d'une morbidité importante mais d'une faible mortalité. Ainsi, son importance économique est-elle considérable notamment en termes de perte de productivité liée aux arrêts maladie. Ainsi, des études américaines intégrant à la fois les coûts médicaux et les pertes de productivité estiment le coût annuel lié aux infections à salmonelles non typhiques entre 500 millions et 3,5 milliards de dollars par an (Frenzen, Riggs et al. 1999). L'étude de Frenzen base ses estimations sur un coût moyen, de 0,5 à 3,8 millions de dollars par décès (selon la méthode d'évaluation

utilisée), de 5 460 dollars par patient hospitalisé, de 315 dollars par patient ayant consulté et de 24 dollars par malade ayant spontanément guéri (Frenzen, Riggs et al. 1999).

En Australie, chaque année, 1,2 millions de visites chez un médecin, 300 000 prescriptions d'antibiotiques et 2,1 millions de jours de congé maladie sont associés aux TIAC, dont les salmonelloses, pour un coût annuel de 1,2 milliards de dollars (Kirk, McKay et al. 2008). En Europe, les coûts annuels liés aux salmonelloses d'origine alimentaire sont estimés entre 560 millions et 2,8 milliards d'euros. Ces estimations sont basées sur un coût par cas de 24 euros à 3,8 millions pour un décès (Anonymous 2001).

Un dernier point peut être abordé ici : le bilan coût bénéfice de la lutte contre les salmonelles. En effet, certains pays ont mis en place des mesures de lutte contre les salmonelles dès l'élevage afin de diminuer le nombre de cas de salmonelloses. Le Danemark, par exemple, estime avoir économisé en 2001 25,5 millions d'euros grâce à un plan de lutte dont les coûts (8,5 millions d'euros) sont intégrés par les professionnels. Les coûts économiques liés aux salmonelloses évalués à 15,5 millions en 2001 auraient en effet été de 41 millions d'euros en l'absence d'intervention (Wegener, Hald et al. 2003).

2.3 Enjeux commerciaux

Les salmonelles sont la première cause de signalement de contamination par des micro-organismes pathogènes dans les aliments au niveau européen en 2008 (Anonymous 2009b). Or, en raison de l'obligation réglementaire d'absence de salmonelles dans certains aliments, dont les produits d'origine animale (Règlement 2073/2005/EC), la détection d'une contamination se traduit par le retrait de l'aliment incriminé.

Par ailleurs, les critères de contamination microbiologique peuvent donner lieu à des restrictions concernant les importations. Ainsi, la Commission Européenne a mis en place une politique de lutte contre les agents zoonotiques, dont *Salmonella*, qui lui permettra d'exiger des pays tiers qu'ils fournissent des denrées présentant des garanties sanitaires équivalentes à celles exigées dans l'Union Européenne (Beloeil 2007). La Suède, précurseur en matière de lutte contre les salmonelles en élevage, bénéficie d'ores et déjà d'un statut « *Salmonella* free » qui lui permet d'imposer des tests supplémentaires sur les denrées d'origine animale importées. Enfin, le risque alimentaire est parfois mis en lien avec des enjeux identitaires, comme c'est le cas au Danemark (Delavigne 2001), ce qui est à l'origine de réflexes protectionnistes. D'une part, des interventions de grande ampleur et dans plusieurs filières (poule pondeuse, poulet de chair et porc) ont été mises en place afin d'éradiquer le risque

domestique de salmonelles (Wegener, Hald et al. 2003). D'autre part, des publications scientifiques issues du Danemark, mais aussi de pays scandinaves, pointent le danger *Salmonella* lié aux voyages ou encore aux produits importés (Kapperud, Lassen et al. 1998; Hald, Vose et al. 2004; Sahlstrom, de Jong et al. 2006; de Jong and Ekdahl 2006b; Hald, Lo Fo Wong et al. 2007). Tout comme les communiqués des autorités nationales, elles alimentent le rejet des produits étrangers (Delavigne 2001) et elles peuvent être utilisées pour justifier la volonté d'appliquer des mesures restrictives aux frontières, justification scientifique exigée par les accords SPS (sanitaires et phytosanitaires) de l'OMC et la Food Law européenne.

Pour un pays comme la France, pays d'élevage et grand producteur de denrées animales, les enjeux commerciaux relatifs à la problématique salmonelles sont importants : image des filières, contraintes à l'export, coût des mesures préventives. En effet, les mesures à mettre en place tout au long de la chaîne alimentaire, de l'élevage à la distribution afin de limiter le risque salmonelles (dépistage, gestion des porteurs, mesures d'hygiène, auto-contrôles...) ont un coût qui se répercute sur le prix de production (Wegener, Hald et al. 2003) et donc sur la compétitivité des filières.

3 La lutte contre les salmonelles

Les salmonelles, comme souligné précédemment, sont des pathogènes zoonotiques ubiquitaires, présents tout au long de la chaîne alimentaire, de l'élevage à l'assiette. De nombreux pays ont mis en place des plans de lutte contre ce pathogène, ciblant divers points du continuum de la chaîne alimentaire (Figure 0.6) avec un certain succès (Edel 1994; Wegener, Hald et al. 2003; Gillespie and Elson 2005; Rostagno, Hurd et al. 2005; Poirier, Watier et al. 2008).

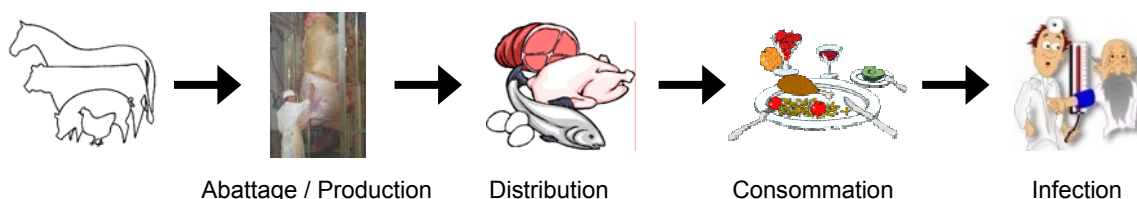


Figure 0.6 : Points d'intervention potentiels pour la lutte contre les salmonelles

3.1 Elevage

Les mesures viseront ici à réduire le niveau de contamination des troupeaux, c'est-à-dire le portage de *Salmonella* par les animaux.

Cela peut reposer sur une démarche d'assainissement des élevages situés en amont des filières organisées, par l'abattage systématique des reproducteurs contaminés, afin de limiter la contamination verticale des filières (Wegener, Hald et al. 2003; Poirier, Watier et al. 2008). Au niveau des troupeaux de production, les mesures prises pourront concerner le contrôle de la contamination des aliments (par acidification par exemple), la vaccination des animaux (Denagamage, O'Connor et al. 2007), le traitement spécifique des troupeaux contaminés (dirigés vers des filières de transformation différentes) éventuellement assorti à un prix indexé sur le niveau de contamination des troupeaux (Wegener, Hald et al. 2003), et enfin une action sur tous les facteurs de risques spécifiques identifiés en production primaire, diminution du stress au cours du transport (Rostagno, Hurd et al. 2005), conduite en bande avec respect de la marche en avant, tout plein tout vide avec nettoyage et désinfection des bâtiments, vides sanitaires entre bandes, etc.

3.2 Abattage et transformation

A ce niveau, la mise en œuvre du système HACCP permet de détecter et contrôler d'éventuelles contaminations en combinaison avec les bonnes pratiques d'hygiène (Edel 1994). En ce qui concerne spécifiquement l'abattoir, agir au niveau de l'ordre d'abattage en traitant les animaux contaminés en dernier peut permettre de limiter les contaminations croisées (Wegener, Hald et al. 2003). Celles-ci sont en effet imputables au transfert, via la chaîne, des contaminants digestifs libérés lors de l'éviscération. La maîtrise de cette étape critique est donc essentielle et peut nécessiter le recours à des dispositifs particuliers de type 'bung bag'¹ (Lo Fo Wong, Hald et al. 2002). De plus, une étape de décontamination des carcasses en fin de chaîne peut se révéler efficace, bien qu'actuellement non autorisée sur le territoire européen (http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753816_1178715010871.htm). On peut noter ici que la traçabilité est un outil indispensable qui permet en cas de contamination d'un aliment de remonter aux matières premières incriminées et au producteur (Sofos 2008).

¹ Le 'bung bag' correspond à l'ensachage de la partie terminale du tube digestif afin d'éviter la dispersion de contenu résiduel dans la cavité abdominale après levée des viscères.

3.3 Distribution

On peut ici évoquer deux points importants, le respect de la chaîne du froid qui permet de limiter le développement des salmonelles en cas de contamination des aliments et de la même façon que pour l'étape précédente, les bonnes pratiques d'hygiène. Enfin, on peut citer l'initiative danoise qui consiste à labelliser les produits issus d'animaux indemnes « *Salmonella free* » (Wegener, Hald et al. 2003).

3.4 Consommation

Enfin, le dernier maillon de la chaîne ne doit pas être négligé. Deux points sont ici essentiels. D'une part, la détection et l'investigation des foyers épidémiques visent à identifier et procéder au retrait de l'aliment responsable des cas de salmonellose groupés (Haeghebaert, Sulem et al. 2003). D'autre part, il est essentiel d'éduquer la population aux bonnes pratiques d'hygiène et de préparation des aliments (Kohl, Rietberg et al. 2002).

3.5 Stratégie de lutte et outils de gestion

Il existe ainsi des mesures de contrôle et de prévention potentielles à chaque point de la chaîne alimentaire. Nombre d'auteurs soulignent l'importance de stratégies intégrées, consistant à agir de façon coordonnée à toutes les étapes de la transmission des salmonelles, de la fourche à la fourchette (Tauxe 1997; Allard 2002; DuPont 2007; Havelaar, Bräunig et al. 2007; Sofos 2008). Cela suppose l'implication de l'ensemble des acteurs, institutions nationales voire supra-nationales (Commission Européenne), professionnels (éleveurs, abatteurs, industrie agro-alimentaire, distributeurs), mais également consommateurs (Allard 2002), et la mise en place d'une réglementation adaptée (Sofos 2008).

Un point essentiel dans la conception d'une stratégie de lutte globale et coordonnée sera de pouvoir identifier les interventions potentielles, de les hiérarchiser en fonction d'un bilan coût-efficacité et enfin de pouvoir évaluer leur impact. Outre la mise en place d'un système de surveillance intégré et efficace, l'un des outils qui permet actuellement de répondre à de tels objectifs est l'attribution de sources. Celle-ci consiste à déterminer la part des infections humaines liées à différentes sources (aliments, animaux, eau, ...) ou différentes voies de transmission (alimentaire, contact, ...) et ainsi de cibler les sources à risque.

4 Méthode d'attribution des salmonelloses humaines aux sources animales

Les salmonelles sont des pathogènes zoonotiques alimentaires majeurs, que l'homme acquière à partir d'une multitude de réservoirs, animaux principalement, et par différentes voies de transmission. Le contrôle et la lutte contre ces infections ont fait l'objet de nombreux efforts, notamment en France, ces dernières décennies. Identifier, prioriser et évaluer les interventions potentielles, est un point clé pour optimiser les stratégies de lutte et décider de l'allocation des moyens publics. Pour ce faire, il est essentiel d'évaluer l'importance relative des différents réservoirs et des différentes voies d'infection, notamment grâce aux méthodes d'attribution. L'article qui suit décrit les différentes approches existantes d'attribution de sources, leur principe, leurs avantages et inconvénients et leur utilité pour répondre aux questions de santé publique. Il s'intitule « Attribuer les cas humains d'infections d'origine alimentaire à des sources spécifiques » et a été rédigé dans le cadre du groupe de travail européen sur l'attribution de source (réseau d'excellence Med-Vet-Net, workpackage 28), cadre dans lequel s'inscrivent également les travaux de thèse. Il a fait l'objet d'une publication dans Foodborne Pathogens and Disease en 2009.

Premier article

Attribuer les cas humains d'infections d'origine alimentaire à
des sources spécifiques

Article publié dans Foodborne Pathogens and Disease

Review

Attributing the Human Disease Burden of Foodborne Infections to Specific Sources

Sara M. Pires,¹ Eric G. Evers,² Wilfrid van Pelt,³ Tracy Ayers,⁴ Elaine Scallan,⁴ Frederick J. Angulo,⁴ Arie Havelaar,^{2,5} Tine Hald,¹ and the Med-Vet-Net Workpackage 28 Working Group⁶

Abstract

Foodborne diseases are an important cause of human illness worldwide. Humans acquire these infections from a variety of sources and routes of transmission. Many efforts have been made in the last decades to prevent and control foodborne diseases, particularly foodborne zoonoses. However, information on the impact of these interventions is limited. To identify and prioritize successful food safety interventions, it is important to attribute the burden of human illness to the specific sources. Defining scientific concepts and harmonizing terminology for “source attribution” is essential for understanding and improving attribution methodologies and for sharing knowledge within the scientific community. We propose harmonized nomenclature, and describe the various approaches for human illness source attribution and their usefulness to address specific public health questions.

Introduction

PATHOGENS COMMONLY TRANSMITTED to humans through foods are responsible for a high burden of human illness and death worldwide. The World Health Organization (WHO) estimates that 1.8 million children die each year from diarrhea, and much of the childhood diarrhea is caused by pathogens that are commonly acquired from contaminated food or water. Furthermore, even in developed countries up to one third of the population each year has an infection from a pathogen commonly transmitted through foods (WHO, 2005). Humans acquire these infections through a number of routes, including eating contaminated food, contact with live animals, and contact with a contaminated environment. Foodborne transmission is recognized as being responsible for a major

proportion of these infections, and foodborne diseases may involve many different food sources and commodities. Several countries have implemented intervention programs during the last decades to prevent and control foodborne diseases, particularly foodborne zoonoses (Wegener *et al.*, 2003; EFSA, 2006). However, precise measurement of the public health impact of such interventions has been difficult, in part because information on the attribution of the burden of foodborne diseases to specific sources is often insufficient. To prioritize appropriate food safety interventions, it is crucial to attribute the human disease burden of each foodborne infection to specific sources (FAO/WHO, 2006).

A variety of general methods to attribute one or more foodborne diseases to specific sources has been developed, including microbiological approaches, epidemiological

¹National Food Institute, Technical University of Denmark, Lyngby, Denmark.

²Laboratory of Zoonoses and Environmental Microbiology, National Institute for Public Health and the Environment, Bilthoven, The Netherlands.

³Epidemiology and Surveillance Unit, National Institute for Public Health and the Environment, Bilthoven, The Netherlands.

⁴Enteric Disease Epidemiology Branch, Division of Foodborne, Bacterial, and Mycotic Diseases, National Center for Zoonotic, Vectorborne, and Enteric Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia.

⁵Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands.

⁶In addition to the authors: Andreas Schroeter (Federal Institute for Risk Assessment [BfR], Germany), Anne Brisabois (French Food Safety Agency [AFSSA]), Anne Thebault (AFSSA), Annemarie Käsbohrer (BfR), Carl Schroeder (United States Department of Agriculture [USDA]), Christina Frank (Robert Koch Institut [RKI], Germany), Chuanfa Guo (USDA–Food Safety and Inspection Service [FSIS]), Danilo Lo Fo Wong (National Food Institute [NFI], Technical University of Denmark), Dörte Döpfer (Wageningen University, The Netherlands), Emma Snary (Veterinary Laboratories Agency [VLA], UK), Gordon Nichols (Health Protection Agency, UK), Heidi Spitznagel (BfR), Helene Wahlström (National Veterinary Institute, Sweden), Julie David (AFSSA), Katarzyna Pancer (National Public Health Institute, Poland), Klaus Stark (RKI), Lars Plym Forshell (National Food Administration, Sweden), Padraig Nally (VLA), Pascal Sanders (AFSSA), and Petra Hiller (BfR).

approaches, intervention studies, and expert elicitation approaches. Each of these general methods presents advantages and limitations, and the usefulness of each depends on the public health questions being addressed (Batz *et al.*, 2005). Several groups are using attribution methods, but these often have different nomenclature and food categorization schemes. Defining scientific concepts and harmonizing terminology are essential for understanding and improving attribution methodologies and sharing knowledge across the scientific community. In this paper, we propose harmonized nomenclature; describe different approaches to human illness attribution; and discuss the advantages, limitations, and applicability of each approach in answering different questions along the farm-to-consumption continuum, while emphasizing that the choice of the attribution method will depend on the pathogen and the public health question being addressed.

Definitions

Human illness “source attribution” may be defined as the partitioning of the human disease burden of one or more foodborne infections to specific sources, where the term *source* includes animal reservoirs and vehicles (e.g., foods). The human disease burden can be measured by the number of laboratory-confirmed (reported) infections or by the estimated total number of infections. To enable comparisons and to account for morbidity and mortality, the human disease burden can be expressed as disability adjusted life years (DALYs).

Attribution of human illness to specific sources requires categorization of the sources. Harmonization of the categorization scheme is needed for comparisons and integration of results from various models and approaches. Such a system should be hierarchical, while accommodating different levels of detail required for different purposes. The categorization scheme should fit with food consumption databases and be internationally standardized.

General Approaches for Source Attribution

Microbiological approaches

One general method for attribution of the human disease burden of foodborne infections to specific sources is “microbiological approaches.” Microbiological approaches for source attribution include the microbial subtyping approach and the comparative exposure assessment approach. Both approaches involve isolation of the pathogen from the various sources and from ill humans. The microbial subtyping approach requires a representative distribution of the subtypes of the pathogen in the different sources and humans, but does not depend on estimates of the prevalence of the subtypes in each source. The comparative exposure assessment approach requires estimates of the prevalence and concentration of the pathogen in each of the sources of exposure.

Microbial subtyping approach. The microbial subtyping approach involves characterization of isolates of a specific pathogen by phenotypic and/or genotypic subtyping methods (e.g., serotyping, phage typing, antimicrobial susceptibility testing, pulsed-field gel electrophoresis, sequence-based subtyping). The principle is to compare the subtypes of isolates from different sources (e.g., animals, food) with those isolated from humans. The microbial subtyping approach is

enabled by the identification of strong associations between some of the dominant subtypes and a specific reservoir or source, providing a heterogeneous distribution of subtypes among the sources. As a first step, subtypes isolated exclusively or almost exclusively from one source are regarded as “indicator subtypes,” and the human infections caused by each indicator subtype are assigned (attributed) to that specific source. The relationship between the relative occurrence (i.e., proportion of positive samples or positive isolates) of each indicator subtype in the source and the incidence of human infections caused by that indicator subtype is then determined. Finally, human infections caused by subtypes found in several sources are assigned to specific sources proportional to the occurrence of the indicator subtypes. The application of this approach assumes that the distribution of subtypes in the collection of isolates in each source used in the attribution exercise is similar to the true distribution of subtypes in each source. Because the microbial subtyping approach utilizes a collection of temporally and spatially related isolates from various sources, it is facilitated by an integrated foodborne disease surveillance program that is focused on the collection of isolates from the major food animal reservoirs of foodborne diseases.

There have been several applications of the microbial subtyping approach for *Salmonella* source attribution (e.g., Van Pelt *et al.*, 1999; Sarwari *et al.*, 2001). The most advanced application of the microbial subtyping approach for *Salmonella* was developed in Denmark (Hald *et al.*, 2004). Using data from the integrated Danish *Salmonella* surveillance program, a mathematical model was developed to quantify the contribution of each of the major food animal sources to human *Salmonella* infections. The “Danish *Salmonella* source account” model attributes domestically acquired laboratory-confirmed human *Salmonella* infections caused by different *Salmonella* subtypes (serotypes and phage types) as a function of the prevalence of these subtypes in animal and food sources and the amount of each food source consumed, using a Bayesian framework with Markov chain Monte Carlo simulation (Gilks *et al.*, 1996). This microbial subtyping approach has proved to be a valuable tool in focusing food safety interventions to the appropriate animal reservoir in Denmark and provides an example of potential synergy between quantitative risk assessment and public health surveillance (Hald *et al.*, 2004).

Another example of the microbial subtyping approach is the use of multilocus sequence typing (MLST) of *Campylobacter jejuni* isolates from foods and humans, being applied in the United Kingdom (Dingle *et al.*, 2002) and New Zealand (French, 2007). In this microbial subtyping approach, MLST is used to identify lineages in bacterial populations by indexing the variation present in seven housekeeping genes located in various parts of the chromosome (Dingle *et al.*, 2001). With the development of novel analysis tools such as the ClonalFrame algorithm (Didelot and Falush, 2007), MLST has been used to identify clonal complexes associated with different isolation sources that, in some instances, correspond to different host species. As an increasing number of isolates, including isolates from various sources, are added to the MLST database (accessible on the internet), there will be increased precision in the attribution of human infections to host sources (McCarthy *et al.*, 2007). Recently, MLST data were utilized to attribute the sources of human *C. jejuni* infections in New Zealand using two microbial subtyping models, the ClonalFrame algorithm

and the Danish *Salmonella* source account; both models gave similar results (French, 2007).

When the microbial subtyping approach is applied at the point of production, it quantifies the contribution of the most important reservoirs of the specific pathogen, assuming isolates are available from those reservoirs. Because pathogens can be transmitted through a variety of sources, interventions that control the pathogen at the reservoir level, before the dissemination of the pathogen through numerous transmission pathways, should result in important declines in human infections.

A limitation is that the microbial subtyping approach is restricted to pathogens that are heterogeneously distributed among the sources. Furthermore, the microbial subtyping approach will only attribute illness to those sources from which isolates are available. The microbial subtyping approach is also data intensive and requires a sufficiently large and representative sample from each source. Independent of the typing method used, the application of strain typing to make inferences on the sources of human infections must be done with appreciation of the strengths and limitations of typing methods (Olsen *et al.*, 1993; Tenover *et al.*, 1997). Clonality is defined as the high probability that two isolates that are identical by typing are related to each other, and the confidence in this probability becomes greater when more than one typing method is applied (Olsen *et al.*, 1993).

Comparative exposure assessment approach. The principle of the comparative exposure assessment approach is to determine the relative importance of the known transmission routes by estimating the human exposure to that pathogen via each route. The comparative exposure assessment approach requires, for each known transmission route, information on the prevalence and dose of the pathogen in the source, the changes of the prevalence and quantity of the pathogen throughout the transmission chain, and the frequency at which humans are exposed by that route. These data provide an estimate of the exposure dose for each transmission route. The exposure doses are compared and the human disease burden (e.g., the observed laboratory-confirmed infections or estimated total number of infections) caused by the specific pathogen is partitioned to each of the various transmission routes, proportionally to the size of the exposure dose. The estimates of exposure dose for each transmission route can be combined with a dose-response model to predict the number of infections from each route, similar to what is done in traditional microbial risk assessments. The comparative exposure assessment approach for source attribution differs from traditional risk assessment in its objective and level of detail. A risk assessment typically aims at describing the complex dynamics of a pathogen in a single food commodity during food processing, and predicting the relative public-health effect of different interventions strategies—alone and in combination. In contrast, the comparative exposure assessment aims at partitioning the observed (or predicted) human disease burden to all known transmission routes, including various foods, direct contact with live animals, and environmental exposures. For this purpose, the various transmission routes are modeled in a more simplified and less detailed way that represents only the main steps in the transmission pathway.

A comparative exposure assessment approach was used to attribute human *Campylobacter* infections to various trans-

mission routes in the Netherlands (Evers *et al.*, 2008). Human exposure across the major transmission routes was estimated in a stochastic model as the mean dose of *Campylobacter* ingested per person per day, averaged over the entire Dutch population. Thirty-one transmission routes related to ingestion of food and direct contact with animals and water were investigated. Another example of this approach was developed to attribute human *Campylobacter* infections in New Zealand, where a model was used to explore the relative importance of four of the most commonly identified routes of exposure: food (poultry and red meat), drinking water, freshwater swimming, and occupational contact (livestock). The model estimates the mean daily dose to which an average individual is exposed to from each route, and, by comparing the relative exposure doses, human *Campylobacter* infections were attributed to the sources (McBride *et al.*, 2005).

The strength of the comparative exposure assessment approach is that it attributes human illness to sources of exposure, taking into account the different transmission routes from the same reservoir, e.g., estimating the role of dairy products and beef from the cattle reservoir. However, this approach is often limited by a lack of sufficient data (e.g., comparable prevalence data or data on exposure through different routes), which results in large uncertainties around the estimates.

Epidemiological approaches

A second general method for attribution of the human disease burden of foodborne diseases to specific sources is “epidemiological approaches.” Epidemiological approaches for source attribution usually involve interviews of patients to elicit the patient’s recall of foods consumed or other exposures before illness began. Some patient interviews are undertaken as a routine activity under the existing public health surveillance infrastructure. Additional studies may be performed to determine factors associated with apparently sporadic infections or in association with investigations of outbreaks of human infections. An outbreak is defined as the occurrence of two or more cases of a similar illness resulting from the exposure to a common source (Olsen *et al.*, 2000), while sporadic cases represent cases that have not been associated with known outbreaks (Engberg, 2006). In some circumstances, cases classified as sporadic may belong to undetected outbreaks. Identification of possible sources of apparently sporadic infections and outbreaks may be undertaken using analytical epidemiological studies, which involve interviewing persons who are (or will become) ill and persons who are not (or will not become) ill, or case-series studies, which involve interviewing only individuals who are ill.

Studies of sporadic infections. Several types of studies have been performed to identify possible sources of apparently sporadic human infections. Case-control studies are the most commonly used analytical epidemiological studies for identifying possible exposures associated with sporadic infections. To allow sufficient enrollment of patients, case-control studies of sporadic infections are often conducted over an extended period of time and commonly use public health surveillance to ascertain culture-confirmed cases. Typically, selected case-patients and a corresponding group of asymptomatic, and therefore assumed to be uninfected, individuals

(controls) are interviewed, and the relative role of exposures is estimated by comparing the frequency of exposures among cases and controls. When infections are associated with an exposure, the proportion of cases attributed to the exposure can be calculated and is defined epidemiologically as the "population attributable fraction" (PAF) (Clayton and Hills, 1993). The PAFs can be used to partition the human disease burden to specific sources (Stafford *et al.*, 2008).

Numerous case-control studies of sporadic infections of diseases commonly transmitted through food, including zoonotic diseases, have been published. Some of these studies calculate only measures of association, e.g., *Salmonella* Enteritidis infections in Denmark (Mølbak and Neimann, 2002) and *Campylobacter* infections in Denmark (Wingstrand *et al.*, 2006), while others estimate the PAF, e.g., Shiga toxin-producing *Escherichia coli* O157 infections in the United States (Voetsch *et al.*, 2007). Case-control studies, when combined with data on the human health burden of the disease under study, can also be utilized to attribute that burden to specific sources. A good example is the recently published study for sporadic *Campylobacter* infections in Australia (Stafford *et al.*, 2008), where the authors determine the PAF for various exposures, and combine the PAF estimates with estimates of the burden of *Campylobacter* to partition the burden to specific sources.

Case-control studies are a valuable tool to identify potential risk factors for human infections, including sources and predisposing, behavioral or seasonal factors (Engberg, 2006). Moreover, in addition to individual case-control studies, a systematic review of published case-control studies of sporadic infections of a given pathogen can provide an overview of the relevant exposures and risk factors for that infection, and a summary of the estimated population attributable fractions for each exposure. An overall population attributable fraction derived from a meta-analysis or weighted summary of several case-control studies of a certain pathogen can be combined with estimates of the burden of illness caused by that pathogen to estimate the burden of illness attributed to each exposure.

A limitation of epidemiological approaches, particularly case-control studies, is that cases that are not associated with a recognized outbreak reflect a mixture of possible sources of exposure, and it may be difficult to distinguish between these exposures. Another limitation is that the statistical power to determine the importance of common exposures often requires enrollment of many participants. Furthermore, a limitation of case-control studies is the lack of accuracy in the recall of exposures by participants (ill and well participants). This misclassification of exposures commonly leads to an underestimation of the burden of illness attributed to specific exposures (Stafford *et al.*, 2008).

Cohort studies, another type of analytical epidemiological study, are used less often for sporadic infections, since they usually require interviewing more persons than is practical, most of whom are not infected. Examples of cohort studies performed to determine the overall disease burden attributable to specific pathogens include the study of sporadic infections by de Wit *et al.* (2002) and the Infectious Intestinal Disease study in the United Kingdom (Wheeler *et al.*, 1999).

Case-series studies of sporadic infections are commonly conducted, particularly for uncommon diseases that have a well-recognized source of infection, to which persons without the infection are infrequently exposed. Examples of case series studies of sporadic infections and the common source for that

disease include botulism associated with home-canned foods (Sobel *et al.*, 2004), *Vibrio vulnificus* and oysters (Shapiro *et al.*, 1998), and *Salmonella* Typhi infections and foreign travel in the United States (Ackers *et al.*, 2000). When an exposure is uncommon in the general population, case series studies of a sporadic infection can be used to determine the frequency of that exposure among the cases, and that frequency can be considered against the proportion of the burden of illness caused by that disease that is attributed to that specific exposure.

Analysis of data from outbreak investigations. Another epidemiological approach for source attribution involves conducting an analysis of the information available from outbreak investigations. An investigation of an individual outbreak may involve both microbiological and epidemiological data; the epidemiological data may be derived from interviewing only infected persons (i.e., a case-series) or from interviewing both infected and noninfected persons. Typically, the microbiological data in an individual outbreak investigation are used to generate hypotheses about the source of the outbreak or to support the results of an epidemiological investigation. Many outbreak investigations are successful in identifying the specific source for the human infections. By conducting an analysis of data from outbreak investigations, the most common food vehicles involved in outbreaks can be identified. A simple analysis or summary of outbreak investigations is useful for attributing illnesses to foods, but often the implicated food item in an individual outbreak is a "complex" food, containing several food items, many of which could be the specific source of the infection. Several methods have been used to include the information of complex foods involved in outbreaks when conducting an analysis of data from outbreak investigations to attribute human illness to sources.

An example of an analysis of data from outbreak investigations for source attribution was developed in the United Kingdom. This analysis used data from individual outbreak investigations reported through national surveillance and population-based studies to estimate the number of human foodborne infections in England and Wales associated with specific food sources (Adak *et al.*, 2005). In this method, individual outbreak investigations that implicated complex foods were not included in the analysis. An alternative method for conducting an analysis of data from outbreak investigations is being developed in the United States. In this method, food items implicated in foodborne outbreaks reported to the Centers for Disease Control and Prevention via the electronic Foodborne Outbreak Reporting System are used to attribute illnesses to specific foods. Food items are categorized into a hierarchical scheme, according to their ingredients. Foods that contain ingredients that are members of a single commodity are considered "simple foods," while foods that contain ingredients that are members of multiple commodities are considered "complex foods." As an example, ground beef is a simple food, whereas meat pie is a complex food. Each implicated food is assigned to one or more mutually exclusive food commodities, according to its ingredients. For outbreaks that have implicated a simple food item, all illnesses are attributed to that single commodity. For outbreaks that have implicated a complex food item, illnesses are partitioned to each commodity in the complex food according to the proportion of illnesses attributed to each of those commodities in outbreaks caused by simple foods. As a result, illnesses in an outbreak due to a

complex food item are attributed to a commodity in the implicated complex food, only if that commodity has been implicated in at least one outbreak due to a simple food. The number of illnesses attributed to each commodity are then summed and used to determine the percentage of disease attributed to each commodity. To estimate the attribution of all foodborne illness in the United States (and not just illnesses involved in reported outbreaks), the percentages of illness for each disease attributed to each commodity were weighted by the estimated annual burden of illness for each pathogen.

A limitation of an analysis of data from individual outbreak investigations is that many outbreaks may not be detected, investigated, or reported. In many cases, a pathogen is not identified, the source is not elucidated, or both. Also, the quality of evidence varies and classification schemes for the data are not consistently used. Large outbreaks, outbreaks associated with point sources, outbreaks that have short incubation periods, and outbreaks that cause serious illness are more likely to be detected, investigated, and reported. As a consequence, the burden of illness in the population attributed to some food sources may be underestimated, which will likely vary by pathogen (Batz *et al.*, 2005). Another limitation is that categorization of foods may differ from one analysis to another. A third limitation of an analysis of data from individual outbreak investigations is that foodborne illnesses included in data from outbreak investigations may not be representative of all foodborne illnesses. Finally, certain food vehicles are more likely to be associated with reported outbreaks than others, which can lead to an overestimation of the proportion of human illnesses attributed to a specific food.

An advantage of data from outbreak investigations is that clear documentation that a specific pathogen was transmitted to humans via a specific food item can be available. Furthermore, an analysis of data from outbreak investigations to attribute foodborne illness to specific sources may sometimes include information on the point of contamination at the farm to consumption chain, particularly when individual outbreak investigations have complete trace back information available. Another advantage is that a wide variety of food vehicles are represented, including less frequently identified food items (e.g., almonds or sprouts). Finally, an analysis of data from outbreak investigations may in some countries or regions be the most readily available source of information for human illness attribution purposes.

Intervention studies

A third general method to attribute the human disease burden of one or more foodborne infections to specific sources is the use of intervention studies. Intervention studies can provide compelling evidence of the burden of illness attributed to a specific source, particularly if the intervention is conducted in a randomized, double-blinded design. Intervention studies have proven useful, for example in developed countries in determining the population disease burden from a particular exposure (e.g., drinking water), and in developing countries in determining the burden of illness from a variety of hygiene-related factors. Intervention studies can be designed as small-scale (e.g., at farm level) or larger-scale (e.g., interventions at a national level) studies to control a certain foodborne disease. Examples of intervention studies include the experiences of reducing the burden of human campylobacteriosis from

poultry meat in Iceland (Stern *et al.*, 2003) and human salmonellosis in Denmark (Wegener *et al.*, 2003).

Accidental interventions, where a decrease or increase in human disease incidence can be observed due to a change in exposure or behavior of the population at risk (e.g., sudden changes in consumption of certain food items) may occur. Examples include the sudden decline in beef consumption due to bovine spongiform encephalopathy in the United Kingdom, and declines in chicken consumption due to outbreaks of avian influenza in Europe. By analyzing changes in consumption and numbers of reported cases of human illness, it is possible to estimate the number of human cases attributable to specific sources. For example, in June 1999, the use of dioxin-contaminated chicken feed in Belgium resulted in the withdrawal of chicken meat and eggs from the market (Vellinga and Van Look, 2002). A decrease in reported human *Campylobacter* infections also occurred in Belgium, beginning in June 1999. Public health surveillance data from preceding years (1994 to 1998) were used to predict the number of human *Campylobacter* infections expected in 1999. The number of reported cases in 1999 was substantially lower (40%) than the expected, suggesting that 40% of reported human *Campylobacter* infections in Belgium could be attributed to chicken (Velling and Van Lock, 2002).

Expert elicitation approaches

A fourth general method to attribute the human disease burden of foodborne diseases to specific sources is "expert elicitation approaches." Expert elicitations are commonly used to address data gaps. Several studies have used expert elicitations to estimate the proportion of human infections that are foodborne, which can be viewed as a first step in source attribution (e.g., Mead *et al.*, 1999; Van Duynhoven *et al.*, 2002; Hoffmann *et al.*, 2007).

One of the approaches used for expert elicitation is the generalized framework for expert judgment studies from heterogeneous expert panels based on the Group Delphi method (Henson, 1997), which has been adapted for source attribution of foodborne illness in New Zealand (Lake *et al.*, 2006). The process involved selection and invitation of experts, distribution of summary sheets for each food-hazard combination, a group meeting and completion of a questionnaire (without discussion). The results of the questionnaires were then aggregated by Monte Carlo simulation of pert distributions with equal weighting of each participant and reported back to the group. A period of discussion followed and the estimates for the questionnaire were then repeated by each expert and re-aggregated. A final report was prepared and sent to each participant with the possibility of further comment. The discussion period was considered valuable in generating consensus and exchanging information (Lake *et al.*, 2006). Another example of the use of a systematic expert elicitation approach for source attribution is the study performed to estimate the fraction of human cases of enterically transmitted illness by five major pathways (food, environment, direct animal contact, human-human transmission, and travel) and by 11 groups within the food pathway. In this study, 16 food safety experts were asked to provide their estimates of the most likely range for each of the parameters, and joint probability distributions were created by probabilistic inversion (Havelaar *et al.*, 2008).

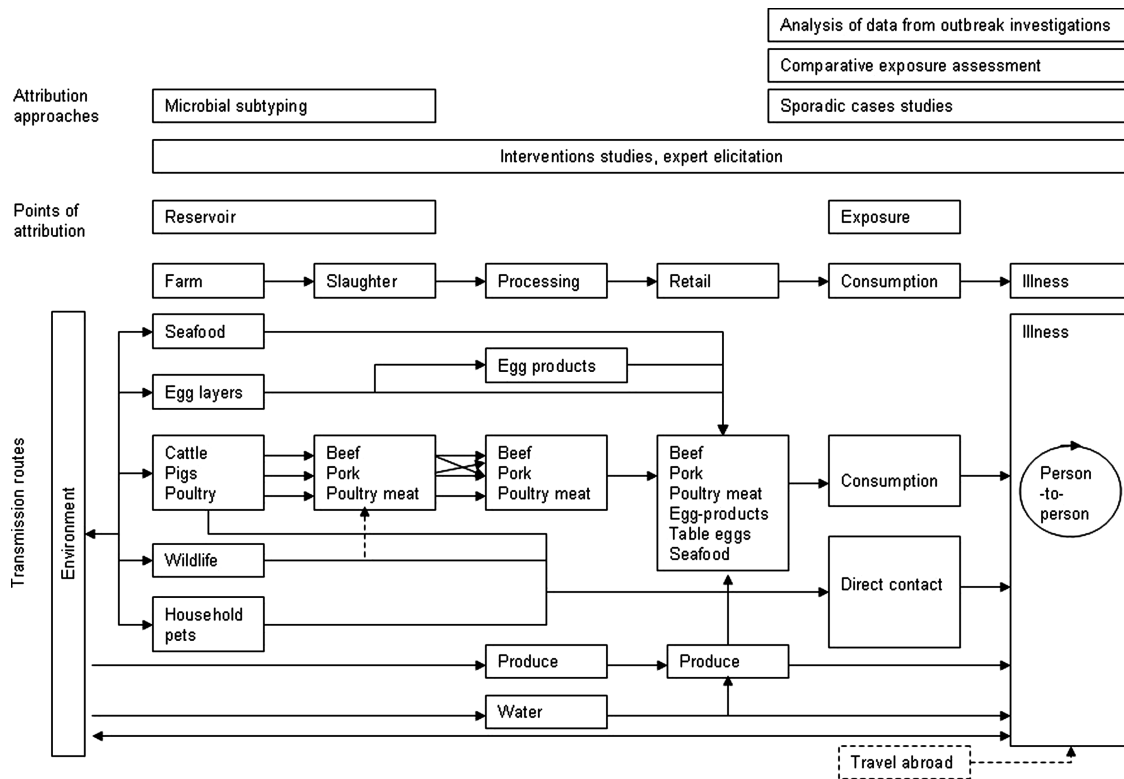


FIG. 1. Routes of transmission of zoonotic pathogens and points of human illness attribution. Travelling abroad is not considered to constitute a route of exposure by itself, as the main routes described will also apply for travelers. Still, it is common to include traveling abroad as a source on its own, which is the reason for including it in this diagram.

Although expert elicitation is useful when data are lacking, a limitation is that the conclusions are based on individual judgment, which may be misinformed or biased (Batz *et al.*, 2004). Nevertheless, for some pathogens or diseases, expert elicitations may be the only available method for source attribution.

Points of Source Attribution

Human illness source attribution can take place at different points along the food chain, including at production, distri-

bution and consumption. Source attribution at the point of production most closely represents attribution of the foodborne pathogen at the reservoir level. A reservoir is defined as an animal species or a nonanimal substance upon which the pathogen depends for its survival (Martin *et al.*, 1987).

Because pathogens that cause foodborne disease may enter the food distribution chain at different points, the burden of illness caused by one disease attributed to specific sources may vary, depending on the point along the food chain the approach focuses on. For example, because cattle are the most important reservoir of *Escherichia coli* O157:H7, attribution of

TABLE 1. HARMONIZED TERMINOLOGY AND METHODOLOGY FOR ATTRIBUTION OF HUMAN ILLNESS TO SPECIFIC SOURCES

Concept	Definitions
Human illness source attribution	Partitioning of the human disease burden of one or more foodborne infections to specific sources
Source	Origin of the pathogen causing infection, including animal reservoirs and vehicles, e.g., foods and water.
Points of attribution	Points in the food chain where human illness attribution can take place, including production, distribution and consumption
Approaches	Methods
Microbiological approaches	Microbial subtyping Comparative exposure assessment
Epidemiological approaches	Studies of sporadic cases Case-control studies Cohort studies Case-series studies
Expert elicitation approaches Intervention studies	Analysis of data from outbreak investigations

E. coli O157:H7 infections will partition more illness to cattle at the point of production (reservoir) than it will partition to beef at the point of consumption, since other foods, besides beef, may be contaminated with the pathogen.

Some of the general methods to attribute foodborne diseases to specific sources work primarily at one point in the food chain (e.g., epidemiological approaches work primarily at the point of consumption), while other general methods (e.g., expert elicitation approaches) can be more generally applied. The general method for source attribution chosen, and consequently the point of attribution, will depend on the availability of data and on the risk management question being addressed. Figure 1 presents the major transmission routes for foodborne infections, including zoonotic infections, and indicates at which point in the transmission chain the different approaches attribute human illness.

Conclusion

To identify and prioritize appropriate food safety interventions and to precisely measure the impact of interventions aimed at controlling foodborne diseases, it is crucial to attribute the human disease burden of foodborne infections to specific sources. Several general methods for human illness source attribution have been developed and are being utilized, and the usefulness of each depends on the pathogen, the specific public health questions being addressed, the data availability, and the strengths and limitations of the different methods.

The limitations of the different general methods for source attribution may be overcome by the integration of different methods. The described approaches often focus on only one point of the transmission chain (e.g., point of reservoir or point of exposure), which means that choosing only one approach for source attribution may be inadequate to answer specific risk management questions. The integration of source attribution approaches aims at improving the estimates at the same point or at different points of attribution. Examples include blending data from studies of sporadic infections and analysis of data from outbreak investigations (integration at the same point of attribution), blending of case-control studies and the microbial subtyping approach, and integration of the microbial subtyping approach and the comparative exposure assessment approach (integration of different points of attribution).

Like epidemiological approaches in general, source attribution approaches describe an association between the outcome (the burden of a given disease) and the specific sources or exposures. Source attribution methods attempt to attribute the burden of disease at the population level, and do not describe causation of disease at the individual level.

Human illness source attribution is increasingly used to partition human illness to the most important sources and as such support risk management strategies. In this paper, we discussed the available source attribution approaches, and introduced nomenclature that should contribute to the harmonization of concepts, definitions, and methods. Table 1 summarizes our proposal for such harmonization. We conclude by encouraging other scientists that apply human illness source attribution to clearly address and define as a minimum: 1) the sources considered, 2) the point(s) of attribution addressed and 3) the attribution approach(es) chosen, because this is crucial for the sharing of knowledge between

research groups and the comparison of results among the scientific community.

Acknowledgments

Med-Vet-Net Work Package 28 was established to focus on the development and comparison of human illness attribution methods. Med-Vet-Net is a European Union "Network of Excellence" of 16 institutions in 10 countries, established to facilitate research for prevention and control of zoonoses and foodborne diseases. The kick-off meeting of the work-package joined members of the partner institutes and participants from the Centers for Disease Control and Prevention and the United States Department of Agriculture Food Safety and Inspection Service, and focused on the harmonization of terminology on attribution. A thorough discussion achieved agreement on important concepts, which have been presented here. More information about Med-Vet-Net and the work package can be found at <http://www.medvetnet.org/cms/>.

Disclosure Statement

No competing financial interests exist.

References

- Ackers ML, Puhr ND, Tauxe RV, *et al.* Laboratory-based surveillance of *Salmonella* serotype Typhi infections in the United States. *JAMA* 2000;283:2668–2673.
- Adak GK, Meakins SM, Yip H, *et al.* Disease risks from foods, England and Wales, 1996–2000. *Emerg Infect Dis* 2005;11:365–372.
- Batz MB, Doyle MP, Morris JG, *et al.* Attributing illness to food. *Emerg Infect Dis* 2005;11:993–999.
- Clayton D and Hills M. *Statistical Models in Epidemiology*. Oxford: Oxford University Press, 1993.
- de Wit M, Koopmans MPG, Kortbeek LM, *et al.* Gastroenteritis in sentinel general practices, the Netherlands. *Emerg Infect Dis* 2002;1:82–91.
- Didelot X, Falush D. Inference of bacterial microevolution using multilocus sequence data. *Genetics* 2007;175:1251–1266.
- Dingle KE, Colles FM, Ure R, *et al.* Molecular characterization of *Campylobacter jejuni* clones: a rational basis for epidemiological investigations. *Emerg Infect Dis* 2002;8:949–955.
- Dingle KE, Colles FM, Wareing DRA, *et al.* Multilocus sequence typing system for *Campylobacter jejuni*. *J Clin Microbiol* 2001;39:14–23.
- [EFSA] European Food Safety Authority. The Community Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents, Antimicrobial Resistance and Foodborne Outbreaks in The European Union in 2005. *The EFSA Journal* 2006;94:288.
- Engberg J. Contributions to the epidemiology of *Campylobacter* infections, a review of clinical and microbiological studies. København: Statens Serum Institute, Lægeforeningens Forlag, 2006.
- Evers EG, Van Der Fels-Klerx HJ, Nauta MJ, *et al.* *Campylobacter* source attribution by exposure assessment. *Int J Risk Assess Manag* 2008;8:174–190.
- [FAO/WHO] Food and Agricultural Organization/World Health Organization. The use of microbiological risk assessment outputs to develop practical risk management strategies: metrics to improve food safety. Available from <http://www.who.int/foodsafety/micro/jemra/meetings/2005/en/>. 2006.
- French N. *Campylobacter*: diversity, host specificity and population biology. *Zoonoses and Public Health*, 54: Abstract 0028.

- 14th International Workshop on *Campylobacter*, *Helicobacter* and Related Organisms, Rotterdam, The Netherlands, September 2–5, 2007.
- Gilks WR, Richardson S, and Spiegelhalter DJ. *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall, 1996.
- Hald T, Vose D, Wegener HC, *et al.* A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Anal.* 2004;24:251–265.
- Havelaar AH, Galindo AV, Kurowicka D, *et al.* Attribution of foodborne pathogens using structured expert elicitation. *Foodborne Pathog Dis* 2008;5:649–659.
- Henson S. Estimating the incidence of food-borne *Salmonella* and the effectiveness of alternative control measures using the Delphi method. *Int J Food Microbiol* 1997;35:195–204.
- Hoffmann S, Fischbeck P, Krupnick A, *et al.* Using expert elicitation to link foodborne illnesses in the United States to foods. *J Food Prot* 2007;70:1220–1229.
- Lake R, Cressey P, and Gallagher E. Expert opinion as a method for estimating attribution of foodborne infection sources. In: Abstracts of the Priority Setting of Foodborne and Zoonotic pathogens, Berlin, Germany, July 19–21, 2006.
- Martin SW, Meek AH, and Willeberg P. *Veterinary Epidemiology: Principles and Methods*. Ames: Iowa State University Press, 1987.
- McBride G, Meleason M, Skelly C, *et al.* Preliminary relative risk assessment for *Campylobacter* exposure in New Zealand: 1. National model for four potential human exposure routes; 2. Farm environmental model. NIWA Client Report: HAM2005-094, 2005.
- McCarthy ND, Colles FM, Dingle KE, *et al.* Host-associated genetic import in *Campylobacter jejuni*. *Emerg Infect Dis* 2007;13:267–272.
- Mead S, Slutsker L, Dietz V, *et al.* Food-related illness and death in the United States. *Emerg Infect Dis* 1999;5:607–625.
- Mølbak K, Neimann J. Risk factors for sporadic infection with *Salmonella* Enteritidis, Denmark 1997–1999. *Am J Epidemiol* 2002;156:654–661.
- Olsen JE, Brown DJ, Skov MN, *et al.* Bacterial typing methods suitable for epidemiological analysis. Applications in investigations of salmonellosis among livestock. *Vet Q* 1993;15:125–135.
- Olsen SJ, MacKinon LC, Gouldin JS, *et al.* Surveillance for foodborne disease outbreaks—United States, 1993–1997. *MMWR CDC Surveill Summ* 2000;49:1–51.
- Reingold AL. Outbreak investigations—a Perspective. *Emerg Infect Dis* 1998;4:21–27.
- Sarwari AR, Magder LS, Levine P, *et al.* Serotype distribution of *Salmonella* isolates from food animals after slaughter differs from that of isolates found in humans. *J Infect Dis* 2001;183:1295–1299.
- Shapiro RL, Altekruze S, Hutwagner L, *et al.* The role of Golf Coast oysters harvested in warmer months in *Vibrio vulnificus* infections in the United States, 1988–1996. *J Infect Dis* 1998;178:752–759.
- Sobel J, Tucker N, Sulka A, *et al.* Foodborne botulism in the United States, 1990–2000. *Emerg Infect Dis* 2004;10:1606–1611.
- Stafford RJ, Schluter PJ, Wilson AJ, *et al.* Population-attributable risk estimates for risk factors associated with *Campylobacter* infection, Australia. *Emerg Infect Dis* 2008;14:895–901.
- Stern NJ, Hiett KL, Alfredsson GA, *et al.* *Campylobacter* spp. in Icelandic poultry operations and human disease. *Epidemiol Infect* 2003;130:23–32.
- Tenover FC, Arbeit RD, and Goering RV. How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. *Infect Control Hosp Epidemiol* 1997;18:426–439.
- Van Duynhoven YTHP, De Wit MAS, Kortbeek LM, *et al.* Voedselinfecties in Nederland. *Ned Tijdschr Med Microbiol* 2002;10:79–83.
- Van Pelt W, Van De Giessen AW, Van Leeuwen WJ, *et al.* Oorsprong, omvang en kosten van humane salmonellose. Deel 1. Oorsprong van humane salmonellose met betrekking tot varken, rund, kip, ei en overige bronnen. *Infectieziekten Bulletin* 1999;10:240–243.
- Vellinga A and Van Loock F. The dioxin crisis as an experiment to determine poultry-related *Campylobacter* enteritis. *Emerg Infect Dis* 2002;8:19–22.
- Voetsch AC, Kennedy MH, Keene WE, *et al.* Risk factors for sporadic Shiga toxin-producing *Escherichia coli* O157 infections in FoodNet sites, 1999–2000. *Epidemiol Infect* 2007;135:993–1000.
- Wegener HC, Hald T, Lo Fo Wong DM, *et al.* *Salmonella* control programs in Denmark. *Emerg Infect Dis* 2003;9:774–780.
- Wheeler JG, Sethi D, Cowden JM, *et al.* Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *Br Med J* 1999;318:1046–1050.
- [WHO] World Health Organization. *The World Health Report 2005—making every mother and child count*. Geneva, Switzerland: WHO, 2005.
- Wingstrand A, Neimann J, Engberg J, *et al.* Fresh chicken as main risk factor for campylobacteriosis, Denmark. *Emerg Infect Dis* 2006;12:280–285.

Address reprint requests to:
 Sara Monteiro Pires, D.V.M.
 National Food Institute
 Technical University of Denmark
 Mørkhøj Bygade 19
 DK-2860 Søborg
 Denmark
 E-mail: smpi@food.dtu.dk

Depuis la rédaction de cet article, d'autres travaux ont été publiés sur le sujet. Deux publications concernent la méthode d'attribution par typage microbiologique (Wilson, Gabriel et al. 2008; Mullner, Jones et al. 2009) et un autre, l'analyse de données issues d'investigation de foyers (Greig and Ravel 2009). Les travaux de Wilson *et al* (2008), portent sur *Campylobacter* et utilisent une méthode de typage génotypique basée sur la variabilité de la séquence de certains gènes (multilocus sequence typing). Alors que le modèle développé par Hald et al (2004) repose sur la prédiction du nombre de cas par source, celui-ci repose sur la distribution des cas entre les sources, mais sans tenir compte des niveaux de consommation, ni des différences entre source et entre types. Les travaux de Mullner *et al* (2009) proposent des variations au modèle de Hald *et al*, afin de traiter le problème de sa surparamétrisation structurelle et de l'adapter à des données de moindre qualité que celles utilisées par les Danois. Le troisième article est un exemple d'attribution à grande échelle. En effet, les données analysées concernent plusieurs pathogènes et une grande variété d'aliments. Elles sont issues de rapports et de publications internationales. Cette étude souligne les différences dans l'importance des sources pouvant exister pour certains pathogènes, dont *Salmonella*, entre différentes zones géographiques (Etats-Unis, Europe, Australie et Nouvelle Zélande,...), la limitation qu'est le biais de publication et la difficulté d'établir un lien spécifique entre certains pathogènes et un aliment, difficulté qui peut être liée notamment aux notions de contamination croisée et de contamination environnementale. Ce dernier point montre l'intérêt des méthodes d'attribution s'intéressant au réservoir animal en amont, telle que l'attribution par typage microbiologique.

5 L'attribution par typage microbiologique

5.1 Pourquoi choisir cette méthode

Comme l'indique la figure 0.7, la méthode d'attribution par typage microbiologique est particulièrement adaptée pour identifier les principaux réservoirs d'un pathogène en amont, au point élevage.

En France, le choix a été fait, pour lutter contre les salmonelles, d'intervenir dès l'élevage. A partir des années 80, un plan d'intervention a ciblé les volailles (poule pondeuse puis poulet de chair) et les mesures prises ont permis de diminuer le nombre de cas humains

(Poirier, Watier et al. 2008). Cependant, les travaux d'évaluation de l'impact du plan *Gallus gallus*, basés sur des analyses de séries chronologiques, soulignent l'impossibilité d'attribuer la baisse des cas liés à Typhimurium spécifiquement à ce plan, étant donné que des mesures étaient également appliquées dans la filière bovin lait et que Typhimurium est un sérotype très présent dans cette filière. Il apparaît donc ici essentiel de disposer d'une méthode permettant de mesurer de façon simultanée le nombre de cas lié à chaque filière, comme l'attribution par typage microbiologique. Ainsi, à partir du nombre de cas lié à chaque filière et de son évolution dans le temps, l'impact des changements de niveau de prévalence dans les filières, induits par les interventions, pourra être évalué de façon spécifique. De plus, les méthodes d'intervention permettent d'identifier des points d'intervention potentiels et de scénariser l'efficacité des mesures envisagées (Miller, Liu et al. 2005). Cependant, cette méthode nécessite de disposer d'un grand nombre de données représentatives concernant les cas humains et la contamination des sources potentielles et suppose l'utilisation à grande échelle d'outils de typage discriminants.

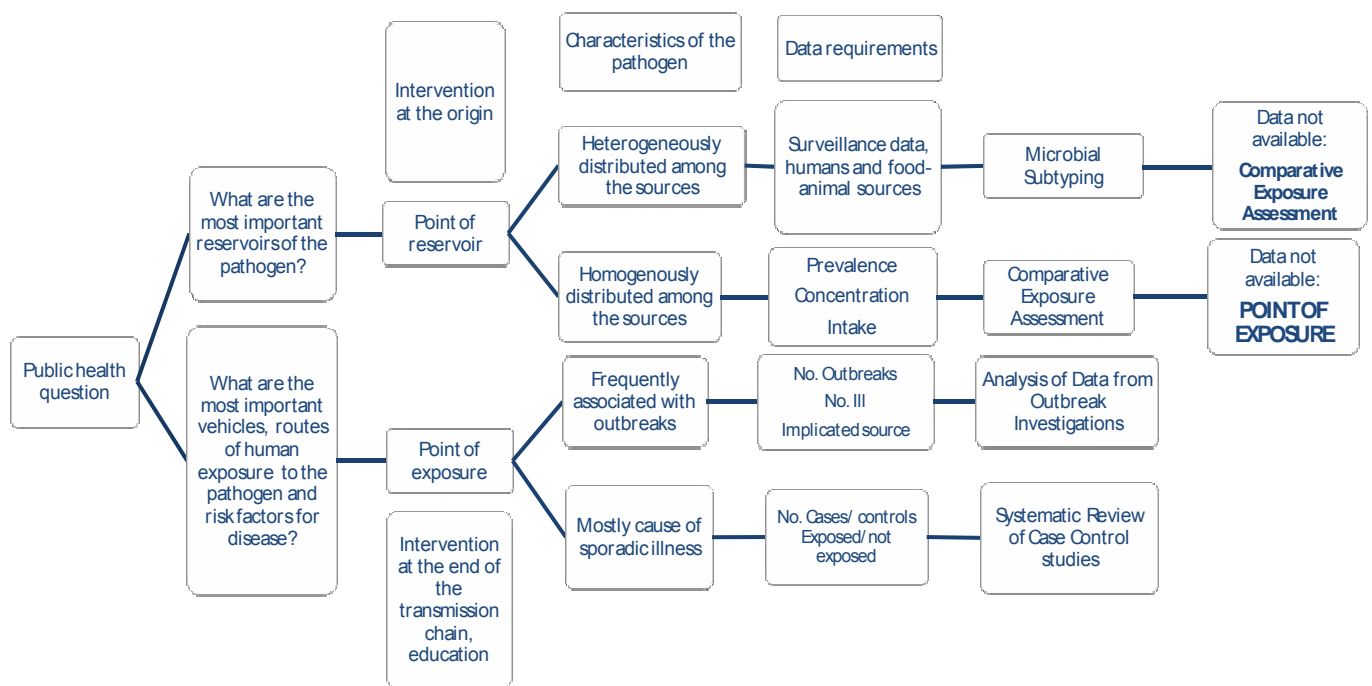


Figure 0.7 : Arbre de décision pour le choix de la méthode utilisée (Rapport final du groupe de travail européen sur l'attribution)

5.2 Principe de la démarche et utilisation du bayésien et de la modélisation bayésienne

La méthode, telle que décrite par Hald et al (2004, Cf § 4.1), repose sur l'estimation du nombre de cas d'un type (de salmonelle) dans une source (λ_{ij}) comme étant proportionnel :

- à la prévalence du type dans la source, ce qui correspond au niveau de contamination (p_{ij});
- à la quantité de source consommée dans la population (M_j) ;
- à la capacité de la source à véhiculer les salmonelles (paramètre source dépendant a_j) ;
- à la capacité du type à causer une infection (paramètre type dépendant q_i).

$$\lambda_{ij} = p_{ij} * M_j * a_j * q_i$$

avec $i=(1, \dots, I)$ et $j=(1, \dots, J)$

Ainsi, l'ensemble des paramètres permettant de cheminer de la contamination des sources observée aux cas humains, notamment en terme de répartition de types microbiologiques est ici pris en compte. Ce nombre de cas attendu est confronté au nombre de cas observé par type, en se basant sur l'hypothèse d'une exposition poissonnienne des cas humains :

$$o_i \sim \text{Poisson} (\sum_j \lambda_{ij})$$

De par son écriture, le modèle est surparamétré (Mullner, Jones et al. 2009). En effet, le nombre d'équations utilisées correspond au nombre de types (I), alors que le nombre de paramètres à estimer (a_j et q_i) est égal au nombre de types plus le nombre de sources ($I+J$). C'est pourquoi il est fait appel au bayésien¹ dans ce cas.

Le problème d'identifiabilité du modèle lié à la surparamétrisation a été traité jusqu'ici de deux façons différentes. Dans l'article originel (Hald, Vose et al. 2004), de l'information « informative » est utilisée en prior pour certains paramètres spécifiques au type. Ainsi pour les sous-types au sein d'un sérotype, les paramètres type spécifiques correspondants sont supposés égaux entre eux. De plus, le paramètre type spécifique du sérotype le plus fréquent parmi les cas humains est utilisé comme référence et fixé à une valeur arbitraire. Aucune hypothèse ni aucune information « informative » ne sont utilisées pour les autres paramètres à estimer à qui sont attribués des distributions *a priori* non informatives (lois uniformes). La solution de Mullner *et al* (2009) est toute autre, une structure hiérarchique

¹ Le principe de la démarche bayésienne est explicité dans l'annexe 1.

est introduite sur le paramètre type dépendant. Les q_i sont alors définis comme des observations aléatoires d'une distribution hypothétique des caractéristiques bactériennes. Cette distribution est définie comme étant une loi lognormale dont la moyenne est constante et fixée et dont le paramètre de précision est à estimer. Le nombre total de paramètres à estimer est donc réduit aux paramètres source-dépendants plus l'hyper-paramètre de la distribution lognormale commune des q_i , soit $J+1$ paramètres.

5.3 Le sous-typage

La méthode d'attribution par typage microbiologique repose sur la caractérisation microbiologique des souches bactériennes. Pour *Salmonella*, la méthode de sérotypage (Cf §1.1.2) est utilisée en première intention. Cependant, deux sérotypes, Enteritidis et Typhimurium, sont toujours isolés très fréquemment parmi les cas humains et sont de plus largement présents dans la plupart des sources. Il est donc nécessaire d'introduire une catégorisation des souches d'Enteritidis et de Typhimurium (Galani, Lo Fo Wong et al. 2006), permettant si possible d'obtenir des sous-types inégalement répartis parmi les sources. La méthode utilisée jusqu'ici dans les travaux utilisant ce modèle d'attribution est la lysotypie (Hald, Vose et al. 2004; Hald, Lo Fo Wong et al. 2007; Pires, Nichols et al. 2008; Mullner, Jones et al. 2009). D'autres méthodes peuvent cependant être utilisées, elles sont soit d'ordre phénotypique (basées sur les caractères biochimiques des souches), soit d'ordre génotypique (Brisabois 2001; Foxman, Zhang et al. 2005; Deplano, Struelens et al. 2007; Hyttia?-Trees, Cooper et al. 2007; Weill 2008; Foley, Lynne et al. 2009).

5.3.1 Méthodes phénotypiques

Les méthodes phénotypiques portent sur l'expression de caractères génétiques parfois complexes. Le phénotype est donc un indicateur imparfait de la variabilité génétique sous-jacente. Dans le cadre d'études visant à déterminer des corrélations écologiques entre des populations de salmonelles issues de différents réservoirs (humains et animaux en l'occurrence), corrélations qui résultent de la parenté de ces populations, ces méthodes seront à utiliser avec précaution.

La lysotypie consiste à étudier la sensibilité des souches à une série de bactériophages sélectionnés. Cette technique est réservée à quelques centres de référence car sa mise en

œuvre, la lecture et l'interprétation sont délicates. Elle est à l'heure actuelle très peu utilisée en France.

L'antibiotypie est basée sur l'étude de la sensibilité des souches à un panel d'antibiotiques. Elle est fréquemment réalisée dans un but thérapeutique, et est également utilisée dans le cadre d'études épidémiologiques (Kariuki, Gilks et al. 1999; Harbottle, White et al. 2006). La résistance aux antibiotiques chez *Salmonella* est déterminée par des gènes situés pour certains sur le chromosome, auquel cas la résistance sera stable pour une lignée, mais également dans une grande partie des cas sur des éléments génétiques mobiles tels les plasmides. Dans ce cas, le profil de résistance sera moins stable et cette méthode sera à considérer avec circonspection quant aux résultats.

5.3.2 Méthodes génotypiques

Les méthodes génétiques visent à détecter le degré de parenté génétique entre différentes souches. Elles sont à ce titre particulièrement adaptées à une utilisation en attribution.

Trois types de méthodes peuvent être mentionnés ici, les méthodes basées sur la restriction de l'ADN, les méthodes basées sur l'amplification de gène et les méthodes basées sur le séquençage. Seules les méthodes les plus fréquemment utilisées sont présentées ici.

Les méthodes de restriction (RFLP, Restriction fragment length polymorphism) reposent sur l'utilisation d'enzymes de restriction, ou endonucléases, qui reconnaissent des sites de coupure dans l'ADN étudié. Leur action mène à la formation de fragments d'autant plus petits que les sites de coupures seront fréquents, fragments qui sont séparés par électrophorèse selon leur poids moléculaire (et donc leur taille).

L'électrophorèse en champ pulsé (PFGE), actuellement considérée comme la méthode de sous-typage moléculaire de référence pour *Salmonella* (des protocoles standardisés ainsi que des bases de données publiques sont développés pour permettre une comparaison de profils rapide et à grande échelle (Swaminathan, Barrett et al. 2001)), fait partie de ces méthodes. Les enzymes utilisées sont des enzymes reconnaissant des sites rares et générant donc un petit nombre de fragments de grande taille. Ces fragments sont séparés par une technique d'électrophorèse utilisant un champ électrique pulsé. La PFGE est très discriminante ; néanmoins, le polymorphisme obtenu après macro-restriction par l'enzyme XbaI (protocole PulseNet) est variable selon le sérotype. Ainsi le pouvoir discriminant, bon pour Typhimurium est relatif pour Enteritidis.

La ribotypie est une méthode RFLP combinant l'utilisation d'enzymes de restriction avec une révélation des fragments d'ADN migrés à l'aide d'une sonde spécifique, sonde ciblant ici les gènes codant pour les ARN ribosomiaux. Le pouvoir discriminant de cette méthode est très variable selon les sérotypes. Néanmoins, elle fait l'objet d'une automatisation permettant standardisation et réduction du temps de manipulation.

Les méthodes d'amplification par PCR (Polymerase chain reaction) sont basées sur l'utilisation de l'amplification de séquences particulières du génome, à l'aide d'une ADN polymérase. Cette méthode permet de détecter la présence ou l'absence de facteurs génétiques particuliers dans le génome de la souche étudiée. Ces méthodes nécessitent de disposer d'amorce correspondant aux facteurs recherchés.

La technique RAPD (Random Amplification Polymorphic DNA¹), basée sur l'utilisation d'une amorce de séquence aléatoire s'hybridant à plusieurs endroits du génome. A la différence de la PCR classique, l'utilisation d'amorces de séquences arbitraires présente l'avantage de ne nécessiter aucune information préalable sur la séquence de l'ADN étudié. Cette technique fréquemment utilisée permet une bonne discrimination au sein d'un sérotype donné. Néanmoins, de par son manque de reproductibilité et de répétabilité, elle n'est pas adaptée à un suivi de souches à long terme.

La MLVA (Multilocus VNTR² Analysis) est basée sur l'examen du polymorphisme de loci comprenant chacun un nombre variable de souches répétées en tandem (VNTR). La méthode consiste à amplifier par PCR ces régions répétées et à analyser la taille des fragments avec un système haute résolution. Il s'agit d'une méthode reproductible et, de plus, particulièrement discriminante pour le sérotype Typhimurium.

Enfin, les méthodes de séquençage reposent sur la détection de variations dans la séquence de certains gènes codant pour les enzymes métaboliques afin de mesurer la distance génétique multi-locus (MLST, Multilocus sequence typing). La grande discrimination du MLST combinée au procédé relativement lent d'accumulation de changements dans la séquence des gènes métaboliques au cours du temps rend cette méthode idéale pour des études épidémiologiques globales.

Une variante de cette méthode est basée sur la détection de polymorphisme pour des nucléotides isolés (SNP, Single nucleotide polymorphisms). Cette dernière option est plus simple mais moins discriminante que la méthode MLST.

¹ Desoxyribonucleic acid

² VNTR : Variable number of tandem repeat

Les méthodes MLST et SNP sont particulièrement adaptées aux études de génotypage à grande échelle. Le principal avantage des méthodes de séquençage pour le typage des souches est l'interprétation non ambiguë des résultats permettant la création de bases de données internationales.

5.3.3 Utilisation en attribution

La méthode de sous-typage idéale pour l'attribution sera une méthode combinant la facilité de mise en œuvre en routine, un bon pouvoir discriminant, une standardisation permettant l'obtention de résultats harmonisés entre laboratoires, et une bonne répétabilité. La méthode doit présenter également des coûts raisonnables afin d'être utilisable à une large échelle. Les caractéristiques des méthodes présentées selon ces critères sont exposées dans la table 0.7.

	Pouvoir discriminant	Reproductibilité	Stabilité	Lisibilité	Coût de mise en œuvre	Exemples d'application à Typhimurium et Enteritidis
Lysotypie	Passable	Passable	Faible	Faible	Elevé	(Boxrud, Pederson-Gulrud et al. 2007) (Guerra, Schrors et al. 2000; Tsen 2002b; Botteldoorn, Herman et al. 2004; Lukinmaa, Nakari et al. 2006)
Antibiotypie	Variable	Bonne	Passable	Excellente	Moyen	(Kariuki, Gilks et al. 1999; Tsen 2002b; Botteldoorn, Herman et al. 2004)
PFGE	Excellent (Selon le nombre d'enzymes utilisé)	Bonne (protocole standardisé Pulse net)	Passable	Bonne	Elevé	(Kariuki, Gilks et al. 1999; Botteldoorn, Herman et al. 2004) (Guerra, Schrors et al. 2000; Bender, Hedberg et al. 2001; Gudmundsdottir, Hardardottir et al. 2003; Fakhr, Nolan et al. 2005; Gebreyes, Altier et al. 2006; Harbottle, White et al. 2006; Lukinmaa, Nakari et al. 2006; Woo and Lee 2006; Boxrud, Pederson-Gulrud et al. 2007)
Ribotypie	Passable	Bonne	Bonne	Bonne	Elevé	(Guerra, Schrors et al. 2000; Jeoffreys, James et al. 2001) (Millemann, Gaubert et al. 2000; Liebana, Garcia-Migura et al. 2002a)
RAPD	Bon	Passable	Passable	Passable	Moyen	(Malorny, Schroeter et al. 2001; Tsen 2002b; Delgado Ronda, Munoz Bellido et al. 2006; Woo and Lee 2006)
MLVA	Bon à excellent	Excellente	Bonne	Excellente	Moyen	(Boxrud, Pederson-Gulrud et al. 2007)
MLST	Bon à excellent	Excellente	Bonne	Excellente	Elevé	(Fakhr, Nolan et al. 2005)
SNP	Excellent	Excellente	Bonne	Excellente	Moyen	

Table 0.7 : Comparaison des performances de systèmes de typage épidémiologique d'après (Foxman, Zhang et al. 2005; Deplano, Struelens et al. 2007)

Problématique

Les salmonelles font partie des agents zoonotiques majeurs dont la maîtrise répond à des objectifs de santé publique. Il s'agit à ce titre d'une zoonose emblématique puisque c'est la première zoonose d'origine alimentaire faisant l'objet d'une approche communautaire intégrée quant à sa surveillance harmonisée et à sa maîtrise dans les filières avicoles et porcines. De plus, alors que production et approvisionnement se globalisent, la France, grand pays producteur de denrées animales, doit faire face à un environnement fortement concurrentiel. Dans ce contexte, les questions de santé publique peuvent constituer des arguments scientifiques forts pour certains pays, afin de limiter les importations ou imposer des contraintes sanitaires sur les denrées importées. Les salmonelles sont au cœur de ce type de préoccupations. Le Danemark, fort d'une démarche volontariste de lutte en filière porcine, a mis en œuvre une démarche d'attribution des salmonelloses humaines aux denrées animales afin de déterminer les contributions relatives de celles-ci et l'allocation des efforts de maîtrise. La France a quant à elle initié très tôt une lutte volontariste contre les salmonelles en filière avicole qui a permis de diminuer le nombre de cas humains. Cependant, les salmonelles demeurent un enjeu de santé publique majeur et dans un contexte de productions animales diversifiées et de rationalisation des efforts des pouvoirs publics, il apparaît crucial pour la France de se positionner sur la démarche d'attribution. Dans le but d'identifier de potentielles interventions au niveau de l'élevage, le choix a été fait de s'approprier l'outil d'attribution par typage microbiologique et d'avoir un regard critique sur cette démarche.

L'objectif de cette thèse a donc été d'évaluer la pertinence de l'outil d'attribution par typage microbiologique en répondant aux deux questions suivantes :

- les données de surveillance produites en France sont-elles suffisantes et adaptées à une telle démarche ?
- le modèle développé par Hald *et al*, est-il robuste et adapté à une utilisation avec des données de surveillance différentes de celles produites au Danemark ?

Les travaux seront ainsi présentés en deux temps. Une première partie est consacrée aux données, sous la forme d'une revue du système de surveillance français puis de l'exposé des

étapes de sélection et d'adaptation des données nécessaires à la démarche d'attribution par typage microbiologique. La deuxième partie concerne plus particulièrement l'outil de modélisation. Dans un premier temps, les essais d'adaptation du modèle en Europe sont présentés au travers d'un projet de publication commun aux participants du projet européen sur l'attribution de sources dans lequel s'inscrit cette thèse. Cet article souligne les difficultés de certains pays dans cet exercice. Nous avons donc cherché à déterminer la raison méthodologique de ces problèmes par une analyse de la sensibilité du modèle à l'information *a priori*, ce qui nous a amenés à proposer une alternative permettant d'obtenir des résultats plus robustes. Enfin, nous nous sommes intéressés à l'impact de la qualité des données sur les résultats d'attribution produits par le modèle à partir des données françaises, dont la particularité est de disposer à la fois de données de surveillance passive et de surveillance active concernant la contamination des sources. Enfin, cela nous permet de dégager les potentialités et évolutions nécessaires du système de surveillance français pour une utilisation des données produites dans une démarche d'attribution, de discuter de l'impact respectif de la méthodologie et de la qualité des données sur les résultats d'attribution et, plus largement, des limites et potentialités de cette démarche.

Travaux personnels

Première partie

Les données

I

Le système de surveillance français des salmonelles

II

Sélection des données et adaptation à l'utilisation en attribution

L'attribution par typage microbiologique est une méthode qui nécessite un volume de données important, couvrant les cas humains comme l'ensemble des sources potentielles de salmonelles. Il est de plus nécessaire de connaître la répartition de types microbiologiques parmi les cas humains et parmi les sources, si possible de façon représentative.

Nous avons donc dans un premier temps décrit et analysé le système de surveillance national des salmonelles, afin de faire le bilan des données disponibles (chapitre I). Ainsi, les bases de données adaptées ou adaptables à une utilisation en attribution ont été identifiées. Ces bases sont actuellement dispersées entre de nombreux acteurs. La collecte et la préparation des données pour ces travaux (chapitre II) reposent donc sur la mise en place d'un réseau de collaborations avec le Centre National de Référence des Salmonelles (CNR), l'Institut de Veille Sanitaire (InVS), la Direction Générale de l'Alimentation (DGAL) et différents laboratoires de l'Agence Française de Sécurité Sanitaire des Aliments (AFSSA). Au-delà de la collecte de données, ce réseau a permis de prospecter l'intérêt de méthodes de sous-typage qui ne sont pas aujourd'hui utilisées en routine et d'engager des réflexions sur différents points méthodologiques.

I

Le système de surveillance français des salmonelles

Deuxième article

Structure du système français de surveillance des *Salmonella* de la ferme à l'assiette

Soumis à Zoonoses and Public Health

Salmonella est un pathogène zoonotique alimentaire qui fait, en France, l'objet d'une volonté de lutte intégrée de longue date, soutenue par le cadre réglementaire européen et s'appuyant sur un système de surveillance exemplaire, couvrant l'ensemble de la chaîne alimentaire, de la fourche à la fourchette. Le système de surveillance intégré des salmonelles en France repose à la fois sur des réseaux de surveillance passive (remontant à 1947 pour les cas humains et à la fin des années 80 pour les principaux réservoirs animaux) et sur un système de surveillance active qui a évolué de façon importante sur les 15 dernières années, sous l'influence de la réglementation européenne.

L'article suivant présente l'architecture du système de surveillance français, ses acteurs et les données produites. Pour chaque secteur, secteur humain puis secteurs vétérinaire et agroalimentaire, le contexte réglementaire européen est d'abord rappelé, puis l'organisation nationale de la surveillance est mise en lumière. La surveillance pour les secteurs vétérinaire et agro-alimentaire a été considérée en trois temps, surveillance des animaux de rente, surveillance des denrées alimentaires et surveillance de l'antibiorésistance. Chaque dispositif, qu'il soit actif ou passif, institutionnel ou non, a fait l'objet d'une analyse détaillée selon les critères de l'OIE (Office International des Epizootie) et de l'EFSA (European Food Safety Agency) concernant l'évaluation des systèmes de surveillance. Enfin, les forces et faiblesse du système, sa dynamique et sa capacité à répondre aux objectifs usuels de la surveillance sont discutées.

Le manuscrit présenté a fait l'objet d'une soumission à Zoonoses and Public Health et est en cours de révision.

Structure of the French farm to table surveillance system for *Salmonella*

J.M. David^{*1,2,3}, C. Danan⁴, C. Chauvin⁵, M. Chazel⁶, R. Souillard⁵, A. Brisabois⁴, FX. Weill⁷, N. Jourdan-Da Silva⁸, M. Picherot⁹, D. Guillemot^{2,3,10}, P. Sanders¹

1 French Food Safety Agency, BP 90203 35302 Fougères, France

2 INSERM, U 657, Paris, F-75015, France

3 Institut Pasteur, Pharmacoépidémiologie et Maladies Infectieuses, 25 rue du Dr Roux 75015 Paris, France

4 French Food Safety Agency, Food quality and agroalimentary processes research Laboratory, 23 Avenue du Général de Gaulle, 94700 Maisons-Alfort, France

5 French Food Safety Agency, Site de Beaucemaine, BP 53, 22440 Ploufragan, France

6 French Food Safety Agency, 31 av Tony Garnier, 69364 Lyon Cedex 07, France

7 Institut Pasteur, CNR des *Salmonella*, Laboratoire des Bactéries Pathogènes Entériques, 28 rue du Dr Roux, 75015 Paris, France

8 Institut de Veille Sanitaire, 12 rue du Val d'Osne, 94415 Saint-Maurice Cedex, France

9 DGAL, Bureau des Zoonoses et de la Microbiologie Alimentaires, Ministère de l'Alimentation, de l'Agriculture et de la Pêche, 251 rue de Vaugirard, 75732 Paris 15, France

10 Univ. Versailles Saint Quentin, Faculté de Médecine Paris Ile de France Ouest, F-78035, France

Corresponding author: Julie David, French Food Safety Agency, Institut Pasteur, BP 90203 35302 Fougères, France. Tel: +33 2 99 94 78 78; Fax: +33 2 99 94 78 99. Email: j.david@afssa.fr

Acknowledgement: This work was funded by the Délégation Générale de l'Armement, Defence Ministry.

Keywords

Food safety, *Salmonella*, Integrated surveillance, Zoonosis, Antimicrobial resistance, Epidemiology

Summary

The French surveillance system is based on a pre-existing national system which can be traced back to 1947 for human cases and the late 80s for the main animal reservoirs. This system evolved under the combined levels of the European regulation and of changes in the observed prevalence of *Salmonella*. European regulation provides strong support to build an active harmonized surveillance system at farm level and for the integration of data throughout the food chain. The parallel existence of passive surveillance networks in the food and veterinary fields allows complementary information to be obtained on other sectors or points of measurement. The main weaknesses and strengths of this system are described and a comparison of the different approaches is presented through a grid analysis. Passive systems are very useful for detecting emerging or unusual events and for early warning of outbreaks. They also produce time series of cases or number of strains that can be used to assess the impact of interventions. Finally, active surveillance data, thanks to their representativeness and reliability, are key elements in the application of risk analysis tools such as quantitative risk assessment or attribution. Thus, despite the dispersal of the data between different actors, the collaboration of these actors at national level and their implication in European and international projects are effective and allow the main objectives of a surveillance system to be addressed.

Bullet points

This review describes the long lasting integrated surveillance system of *Salmonella* in France which evolved in the European regulatory framework and is one of the most performing ones.

It relies on an active surveillance system based on national and European regulations completed by a passive surveillance system composed of several networks.

This system, relying on several different actors, allows treating most of the missions of any surveillance system, from the investigation of outbreaks to complex approaches such as human cases attribution.

Introduction

In the context of the industrialization and the globalization of the food supply, of risk behaviours among consumers (raw eating, undercooking, poor kitchen hygiene) and of the steady increase of susceptible populations such as the elderly and immune-deficient patients, foodborne diseases remain a priority issue despite all the efforts made to prevent them (Sofos 2008).

Human salmonellosis reported in Europe has decreased over the past three years, but *Salmonella* remains one of the most frequently reported causes of foodborne zoonoses (Anonymous 2009). For the period 1995-1999, the French Institute for Public Health (InVS) estimated the total annual number of confirmed cases in France to be between 30,600 and 41,140, of which 92 to 535 led to the death of the patient (InVS 2004).

Moreover, antimicrobial resistance (AMR) has been described in *Salmonella* to antimicrobials used to treat acute gastroenteritis in human (third generation cephalosporins, fluoroquinolones) and has been defined as a risk factor increasing morbidity and mortality in humans (Molbak 2005). It is thus of the utmost importance to control the transmission of *Salmonella*, and consequently that of resistant strains of *Salmonella*, from animals to humans and to monitor antimicrobial resistance levels.

To determine optimal management measures to control this zoonosis, it is necessary to assess the risk of consumer exposure associated with the *Salmonella* contamination of food throughout the food chain. This requires information on prevalence, serotypes, antimicrobial resistance, food consumption, etc.... In this context, monitoring systems for *Salmonella* have been implemented in the human and the agro-food sectors. Such systems must include a trained field force, diagnostic laboratory support, data collection and analysis capabilities (Hueston 1993). To have a global overview of the surveillance systems involved in *Salmonella* monitoring in France and to analyse their specificity and complementarity, an analysis grid has been defined based on OIE (terrestrial animal health code, chapter 6.5) and the European Food Safety Authority (EFSA) guidelines (2006) and has been applied following the usual evaluation steps (CDC 2001). Only current surveillance systems have been included in this analysis.

This paper thus presents the structure, limits and evolution of the French surveillance system relating to *Salmonella* within the framework of the European “farm to table” monitoring approach, focusing on the main reservoirs: food animals and related products.

1 Architecture of the *Salmonella* surveillance system

Salmonella, as a zoonotic pathogen, is the subject of several European regulations. This chapter presents the European regulatory framework and the organization of the national surveillance system, from human cases to animal sources. The actors involved in this system have previously been described by Dufour and La Vieille (Dufour and La Vieille 2000) (Figure 1).

1.1 Human sector

1.1.1 European regulatory framework

The surveillance of sporadic salmonellosis (i.e. not associated with an outbreak) is not regulated at European level. However, surveillance of communicable diseases based on networking of expertise in the EU Member States is promoted by Commission Decision 2119/98/EC. In 2005, under the Zoonosis Directive (Directive 2003/99/EC), the investigation and reporting of foodborne outbreaks became mandatory (Gervelmeyer, Hempen et al. 2008). At the same time, the European Centre for Disease Prevention and Control (ECDC) was established (Regulation (EC) no. 851/2004) in order to enhance the capacity of the Community and the Member states to protect human health through the prevention and control of human diseases. As far as reporting is concerned, EFSA and ECDC jointly analyze all data from the human and agro-food sectors. The results are published in an annual Community Report (2009).

1.1.2 National surveillance system (table 1)

The declaration of foodborne outbreaks has been mandatory in France since 1952 (Decree no. 52-953 of 7 August 1952), allowing outbreak cases surveillance. Foodborne outbreaks are defined as a group of at least two cases with mostly digestive symptoms that can be linked to a common food source. An outbreak will be confirmed as caused by *Salmonella*, if a *Salmonella* strain is isolated from a sample coming either from one of the patients or from the suspected food item. Outbreaks must be notified to either the official “départemental” (“départements” are subdivisions of the national territory administered by a Prefect) sanitary or veterinary services (named DDASS or DDSV respectively). InVS centralizes those data at national level. The DDSV and DDASS databases are merged and

duplicate data are eliminated. In 2007, data on 1,436 foodborne outbreak cases associated with *Salmonella* were centralized. For some outbreaks, specific epidemiological investigations are performed, using data collected during further enquiries and data provided by the National Reference Centre (NRC) and the *Salmonella* Network for the food data (described below in § 1.2.2.2) (Espié, Weill et al. 2005a; Espié, De Valk et al. 2005b; Dominguez, Jourdan-Da Silva et al. 2008). Descriptive results and epidemiological studies are published in the weekly epidemiological bulletin (Bulletin épidémiologique hebdomadaire, <http://www.invs.sante.fr/BEH/>) and in scientific publications (Vaillant and Espie 2005; Brouard, Espié et al. 2007; Jourdan, Le Hello et al. 2008).

No mandatory notification is required for sporadic salmonellosis cases. However, passive laboratory-based surveillance of *Salmonella* detection from human samples, carried out by NRC since 1947, allows the collection of data on sporadic cases. This surveillance system relies on a stable network of voluntary clinical laboratories (private or hospital-based) representing 30 to 40% of all French clinical laboratories involved in human medicine. Laboratories send NRC either strains of *Salmonella* or reports on the strains isolated and serotyped on a daily basis. Strains and reports are registered together with epidemiological information such as travel, age and sex of the patient, type of sample taken (stool, blood...) or geographic location. A biological specimen bank and a national database on human strains are constituted. NRC registers data on about around 10,000 cases a year. Statistical analyses are performed on a weekly basis to detect unusual events, and all surveillance data are communicated to InVS. Data and analysis are published in the annual activity report of the NRC (<http://www.pasteur.fr/sante/clre/cadrecnr/salmcncr/salmcncr-actualites.html>) and through scientific publications (Weill, Guesnier et al. 2006).

1.2 Agro-food sector

1.2.1 European regulatory framework

***Salmonella* in food animals**

The first step in *Salmonella* surveillance in food animals at European level was the implementation of Council Directive 1992/117/EC. This Directive concerned measures for protection against specified zoonoses and zoonotic agents, in animals and products of animal origin, in order to prevent outbreaks of foodborne infections. One goal of Directive 1992/117/EC concerned the control of *Salmonella* in the

Gallus gallus species, i.e. broilers and laying hens, focusing especially on the eradication of the serotypes Enteritidis and Typhimurium in breeding animals.

In 2003, this Directive was replaced by Directive 2003/99/EC of the Council and Parliament and Regulation (EC) no. 2160/2003. Directive 2003/99/EC on the monitoring of zoonoses and zoonotic agents aims to improve and coordinate the monitoring of zoonotic agents in the Community and to collect data that is easier to compile and compare. This should enable hazard identification and characterization, and exposure assessment related to zoonotic agents. *Salmonella* and its antimicrobial resistance are covered by this harmonized monitoring. Regulation (EC) no. 2160/2003 describes the progressive and proportionate organization of the control of *Salmonella* and other specified foodborne zoonotic agents at European level. Member States must implement programs to reduce the prevalence of *Salmonella* in farm animals and products of animal origin. Poultry (broilers, laying hens and turkeys) and pigs, considered as the major animal reservoirs, are the first productions concerned by these control measures at breeding and production level. Within this framework, harmonized prevalence studies, so-called “baseline studies”, have been conducted since 2004, to obtain scientifically relevant data about the initial level of prevalence in each Member State. On the basis of these results, mitigation targets are set species by species for specified serotypes. For example, a maximal prevalence of 1% must be met for the serotypes Hadar, Infantis, Virchow, Enteritidis and Typhimurium by the end of 2009 in breeding flocks of laying hens and broilers (Regulation (EC) no. 1003/2005).

***Salmonella* in food**

In 2002 the European Parliament adopted the “Food Law”, (Regulation (EC) no. 178/2002), laying down the general framework to ensure a coherent approach in the development of food legislation from farm to table; EFSA was created as part of this framework. The “Food Law” establishes principles and responsibilities, the means of providing a strong scientific base, efficient organizational arrangements and procedures to underpin decision-making in matters of food and feed safety. The principle of transparency for the consumer is also established. A package of three regulations and one directive constitutes the food hygiene legislation dedicated to food business operators (FBO) and is completed by two other regulations relating to official controls and feed hygiene.

General rules for FBOs, including primary production, are laid down by Regulation (EC) no. 852/2004. FBOs must ensure that their products satisfy the hygiene requirements set by Regulation (EC) no.

2073/2005 laying down the microbiological criteria and the implementation of hygienic rules. One of the criteria set is the absence of *Salmonella* in the product for safety criteria (i.e. defining the acceptability of the product) and at specific stages of the process for process hygiene criteria (i.e. setting an indicative contamination value above which corrective actions are required). This is to be met according to HACCP (Hazard Analysis Critical Control Points) principles and be scientifically justified. To ensure verification of compliance with the “Food Law”, official controls are performed according to Regulation (EC) no. 882/2004. They must be planned on the basis of risk assessment, relying on a scientific approach, in order to obtain an objective selection of products and operators to be controlled.

Antimicrobial resistance in food animals

Since the end of the nineties, several international scientific reports and recommendations have led to publications recommending the harmonization of surveillance and the regulation of foodborne antimicrobial resistance and the use of antimicrobials in animals, based on public health concerns (WHO, 1997 and 1998; Copenhagen, 1998; FAO/WHO/OIE 2003 and 2007, Codex 2005, OIE 2006). Consequently, a European strategy against the microbial threat has been defined in agreement with the “precautionary principle”. The use of antimicrobials as growth promoters in animal productions has been banned (Council Regulations (EC) no. 2821/98 and (EC) no. 2788/98). At the same time, surveillance systems for non-human use of antimicrobials and for antimicrobial resistance have been implemented at all levels of the food chain and a Community Reference Laboratory has been nominated and financed to promote harmonization of the methods used to assess antimicrobial resistance. For *Salmonella*, a harmonized continuous monitoring system in food animals is being implemented on the basis of a selection of strains isolated during the mandatory control programs (Decision 2007/407/EC). National data on antimicrobial resistance have had to be declared to and have been published by EFSA since 2003 (2009).

1.2.2 National surveillance system

The national surveillance system for *Salmonella* in the agro-food sector relies on active institutional surveillance coordinated by the French Food Directorate (DGAL) operating under the European regulatory framework, completed by a passive system based on networks centralizing non-human

strains of *Salmonella* from public and private veterinary laboratories. Both active and passive surveillance systems are presented following.

Active surveillance system

Salmonella in food animals (table 2)

Based on a pre-existing voluntary program, the continuous mandatory monitoring and control plan in breeding flocks of *Gallus gallus* has been implemented in France since 1999 in accordance with Directive 1992/117/EC. The serotypes Enteritidis and Typhimurium were targeted in breeding flocks of broilers and laying hens until 2007. Since 2007, this plan has evolved in line with Regulation (EC) 2160/2003 and its implementing regulations concerning the control of *Salmonella*. It is in the process of being widened to include turkeys and pigs, all species being monitored at breeding and production level, and is still targeted on the serotypes Enteritidis and Typhimurium. In *Gallus gallus* breeding flocks, 3 more serotypes are regulated: Hadar, Infantis and Virchow. These regulated serotypes are defined as notifiable disease submitted to sanitary policy measures (Rural Code, article D223-21). However, information on all the serotypes will be made available thanks to the classification of all *Salmonella* serotypes as notifiable diseases not subject to sanitary policy measures (Rural Code, article D223-1) for the regulated animal sectors. The continuous monitoring program concerns all flocks and herds and is thus exhaustive for the regulated sectors. This should lead to an annual database giving information on the national prevalence of the different *Salmonella* serotypes and their evolution in the regulated animal species at farm level.

To prepare the enforcement of these monitoring and control programs, the initial level of prevalence of *Salmonella spp* has been evaluated through baseline studies between 2004 and 2009. Sampling plans have been designed to assess a prevalence of 20 to 50 % (according to the species) with a precision of 3 to 5%. They have been conducted under the responsibility of DGAL, by national and departmental veterinary services in collaboration with the National Reference Laboratory (NRL) for *Salmonella*.

Salmonella in food (table 3)

In accordance with the "Food Law", DGAL implements national plans either to assess consumer exposure ("surveillance plans" providing science-based evidence) or to detect anomalies or non conformities by the FBO ("control plans"). In the first case, the sample is randomly defined; in the

second case, the sample is targeted on one type of food operator or one food category, including imported products. If necessary, the French food safety agency (AFSSA), as NRL, is consulted for the definition of the protocols. In accordance with Regulation (EC) no. 882/2004, control and surveillance plans performed by veterinary services are implemented on a yearly basis. The data collected give information on the prevalence of *Salmonella* in different food products, at the slaughterhouse and/or at retail level.

Antimicrobial resistance monitoring

Since 1999 the Food Directorate has supported monitoring plans for antimicrobial resistance in indicator and zoonotic bacteria in the major animal production sectors (pig, poultry and cattle). These plans have been managed by AFSSA in collaboration with the DDSV and public laboratories (LVD) since 1999 for broilers, 2000 for pigs and 2002 for cattle. These plans are still ongoing, but the search for *Salmonella* has been progressively abandoned (in 2004 for broilers and pigs and in 2008 for cattle) because of the low prevalence leading to an insufficient isolation of strains.

Since 2008, surveillance of antimicrobial resistance in *Salmonella* has been implemented in the regulated sectors (*Gallus gallus*, turkeys and pigs), beginning with laying hens in 2008. The strains tested for antimicrobial resistance come from the continuous monitoring and control plans with an upper limit of 170 per year and production. For the food sector, recent data have been collected under the national control and surveillance plans.

This surveillance activity has been included in the working plan of the NRL for antimicrobial resistance following technical recommendations of EFSA for European data harmonization. It should lead to information on antimicrobial resistance in *Salmonella* serotypes per animal species or food category.

Thus, the active institutional surveillance furnishes national data on the prevalence of *Salmonella* per serotype, with the flock or herd as epidemiological unit for food animals and the batch or carcass for food. Depending on the studies, the target populations include breeding animals, production animals or food categories, covering all the food-chain. The different studies are based on randomly chosen samples generally with stratification on the size of the farm or slaughterhouse, designed to be representative of national production for food animals and to target risk items for food. The continuous mandatory monitoring is exhaustive for the targeted productions. Long-term continuous mandatory

monitoring allows trends in prevalence and emergence of particular serotypes or resistance profiles to be analyzed, whereas transversal studies (baseline studies, control plans and surveillance plans) give isolated evaluations of the prevalence. The laboratories involved in this surveillance work under quality assurance and the NRL organizes inter-laboratory assays to ensure the harmonization of the methods. Data are analyzed either by the DGAL or the NRL for the baseline studies. The data collected in the context of harmonized European surveillance are further analyzed and published by EFSA in annual community reports. The data are all published on internal communication media and, for the food data, on the agriculture ministry website.

Passive surveillance system (table 4)

For several years, AFSSA has been managing several passive surveillance networks that provide data on *Salmonella* isolated in several sectors of the agro-food chain.

The epidemiological observatory in poultry farming network (RNOEA) was created in 1987 and is managed by the AFSSA laboratory at Ploufragan (Souillard, Toux et al. 2007). The objective is to give to veterinarians epidemiological information on diseases observed in poultry, to follow trends and detect emergence. This network alerts about major diseases in poultry. The available data for *Salmonella* will concern the distribution of serotypes among both pathogenic and healthy carriage strains found in poultry.

The French antimicrobial resistance surveillance in veterinary pathogens network (RESAPATH) was created in 1982. This network is managed by the coordinated action of two AFSSA laboratories, Lyon (for cattle and small ruminants) and Ploufragan (for poultry and pigs) (Jouy, Meunier et al. 2003; Botrel, Chazel et al. 2006). Its aim is to give scientific and technical advice to laboratories and veterinarians on antimicrobial resistance in pathogenic bacteria, including *Salmonella*, isolated from food-producing animals mostly exhibiting clinical signs. Data on trends and emergence of antimicrobial resistance in pathogenic *Salmonella* are produced by this network.

The "*Salmonella*" Network was created in 1997. This network is managed by the AFSSA laboratory at Maisons-Alfort and collects strains of non-human origin. Its objectives are to provide technical advice for *Salmonella* serotyping, give national temporal trends on *Salmonella* serotypes isolated throughout the food chain and detect emerging or unusual events (Moury, Fremy et al. 2006). Data available will

concern the distribution of serotypes and the antimicrobial resistance profiles among 13 to 14,000 non-human strains of *Salmonella* a year, covering the whole food chain (from animals to products).

All the networks' partners are private and public veterinary laboratories. In addition, veterinarians provide case records to the RNOEA. Microbial analyses are multicentric (partly performed by AFSSA for the *Salmonella* Network); only the results of biological tests or diagnostics are reported to the RNOEA or the Resapath. However, analyses are validated by inter-laboratory trials organized annually by AFSSA, for the *Salmonella* Network (serotyping) and the Resapath (antimicrobial resistance). In all cases epidemiological data are collected on the original sample (geographical data, date of sampling, product description). Duplication of data is not totally under control. For the *Salmonella* Network, antimicrobial resistance testing is performed on "independent" isolates defined as coming from the same parcel, belonging to the same serotype and sharing similar epidemiological data. For the RESAPATH, duplicate data are defined for strains isolated from the same animal or flock, the same date of sampling and sharing the same serotype and antibiotype. Descriptive data are regularly reported to the partners. Detection of unusual events is regularly performed on the number of isolates per serotype at the *Salmonella* Network.

Public networks available for professionals concerned and scientists have been described, but this description does not preclude the existence of other databases in France maintained by and available to private operators on specific topics. For example, specialized private networks focus on *Salmonella* isolation results on feedstuffs (Qualimat network) and on *Salmonella* strains isolated in the pork sector (http://www.ifip.asso.fr/actu/pdf/MA_aval.pdf).

2 Discussion/ conclusions

The surveillance of the agro-food sector in France relies on an integrated food-chain system implemented under the European regulations, completed by lab-based passive networks. European regulation offers strong support for surveillance in the agro-food sector, pushing forward the previously existing systems toward harmonized methodologies, systematic sampling at farm level in the main sectors (Zoonosis Directive), risk-based surveillance for food within the framework of the "Food Law" (Stark, Regula et al. 2006) and integration of antimicrobial resistance in the monitoring systems. The importance of harmonizing the surveillance tools (typing tools, analyses, ...) is emphasized by many

authors (Batz, Doyle et al. 2005; Gerner-Smidt, Hise et al. 2006; Ammon and Tauxe 2007). It relies in Europe on the NRLs, coordinated by the Community Reference Laboratories through their technical support and the inter-laboratory assays they organize (Ellerbroek, Alter et al. 2009). The final objective of this *Salmonella* surveillance system is to contribute to the reduction of the incidence of human cases with the choice of a farm to table approach aimed at limiting contamination at each stage in the food chain. This is of the utmost importance as small improvements in animal health can lead to a significant reduction in human cases according to the model developed by Singer (Singer, Cox Jr. et al. 2007), and as no fail-proof way of assuring the safety of food items exists at any point in the food supply chain (Allard 2002).

Limits and threats

Many of the usual limits and structural threats encountered in health surveillance systems are not avoided in the French *Salmonella* surveillance tools.

The burden of salmonellosis is difficult to assess and is probably underestimated, as the reporting of human diseases in France is based both on mandatory notification of outbreaks and the centralization of passive laboratory results. These surveillance systems, as most foodborne and declarative surveillance systems, are subject to underreporting (Flint, Van Duynhoven et al. 2005) with the cases reported and registered being only a part of the effective burden (Angulo, Voetsch et al. 1998). The mandatory reporting of foodborne outbreaks related to *Salmonella* has been estimated by a capture-recapture method to represent 15% of all cases (Gallay, Vaillant et al. 2000). This may be linked to complex bureaucratic procedures (Herida 2008) and the unknown aetiology of most reported outbreaks (Gallay, Vaillant et al. 2000). Similarly, the extent of *Salmonella* animal carriage or food contamination is not precisely known. Mandatory exhaustive control plans and prevalence studies do not cover all species. Fowls, turkeys, pigs are surveyed at farm and slaughterhouse level, cattle and sheep are concerned by controls performed at slaughterhouse level, but other food animals can contribute to the transmission of *Salmonella* such as fish, ducks (Hald, Vose et al. 2004) or dairy animals and other products (Haeghebaert, Sulem et al. 2003; Lailier, Sanaa et al. 2005). No representative data are available for such sources. For sources concerned by prevalence studies or

an exhaustive mandatory control plan, it can also be noted that small herds are not included and that any investigator bias can not be excluded despite strong harmonization efforts.

Meanwhile, all sources are concerned by passive surveillance, but such systems suffer from a lack of representativeness and reporting delay. The willingness of the laboratories to participate in the networks can interfere, and selection bias can affect the human cases reported or the strains collected. In addition to biases due to the passive structure of networks, their multicentric design can lead to investigator bias. As an example, the epidemiological information associated with the strains is of unequal quality. As a result, the management of duplicates can be difficult. All these weaknesses are highly dependent on the functional quality of the network. The evaluation of such networks is therefore a necessary procedure to determine the reliability of the results (Hendrikx, Chazel et al. 2005).

In France, the actors involved in the integrated surveillance of *Salmonella* are numerous. It makes the centralization of the data difficult and leads to an unequal quality and heterogeneity in the data collected. For example, the epidemiological units and the subtyping tools applied are diverse. It can limit collaborative studies and data cross analysis. Moreover, this “multiple head” system has a non-negligible inertia in a context where there are an obligation to comply with European regulations and a constant evolution of prevalences and exposure. Thus, on the one hand, a few sectors, known for their potential high prevalences and major exposure are well monitored. But thanks to the control programs implemented, their prevalence levels are mitigated (the prevalences of Typhimurium and Enteritidis in *Gallus gallus* breeding flocks are already under 1% in France). On the other hand, other sectors (such as ducks, minor species...) which have a lower, but non-negligible consumption and may present high prevalence rates do not benefit of the active surveillance system.

Strength and opportunities

Some of the characteristics of this surveillance system, despite the inherent biases previously exposed, can also confer strengths on the system.

The multiple part surveillance system, made up of different independent networks and actors, allows a high level of specialization of each one with a clear definition of objectives, a strong involvement of participants and a good knowledge of the data and their characteristics. Each network has a strong

background. The fact that they are long established means that they are able to rely on stable and reliable networks. The networks of laboratories participating in passive surveillance systems such as those of NRC or AFSSA have remained stable over the years and cover all the territory. The resulting databases are therefore appropriate to detect trends or unusual events for both sporadic and outbreak-related cases and food source contamination.

Adaptation of the surveillance system to its subject evolution is also possible with such a flexible organization thanks to several independent tools. For example, the active antimicrobial resistance surveillance system was suspended for *Salmonella* in 2002 for poultry and pigs and more recently for cattle. Too few strains could be collected according to the prevalences reached, whereas on one hand an exhaustive control plan was implemented (in poultry production) and on the other hand a passive network (*Salmonella* Network) could collect more strains. Specialized passive networks can also evolve considering prevalence mitigation as for the RESSAB. The clinical salmonellosis in adult cattle network (RESSAB) created in 1996 was run by AFSSA until 2007. The objectives were to monitor the incidence of salmonellosis associated with digestive or abortive symptoms in adult cattle and to identify associated serotypes and antimicrobial resistance (Chazel, Buret et al. 2007). This network made information available on the prevalence of lab-confirmed clinical salmonellosis in cattle, but ceased its activity in 2007 when the incidence became too low.

In addition, independence of surveillance tools does not exclude close collaboration and harmonization such as between NRC and InVS, and with the *Salmonella* Network, reinforcing the complementarities of their databases (Kerouanton, Marault et al. 2007).

Complementarities are also obvious through the combination of active and passive tools. In a context of limited public resources allocation, active surveillance can not cover all sectors annually. This incompleteness of surveillance is partly addressed by the co-existing permanent passive system. In the same way, the national *Salmonella* surveillance system described here can also collect data belonging to non-food animal associated sources and cases. Official surveillance of the non-animal food sources, such as vegetables (Brandl 2006), is partly performed by the General Directorate for Competition Policy, Consumer Affairs and Fraud Control through annual surveillance and control plans. Non-food sources are also to be considered since an estimated 5% of the transmission of *Salmonella* is not food-mediated (Mead, Slutsker et al. 1999). Pets and especially exotic pets are known to be contamination sources (Woodward, Khakhria et al. 1997; Mermin, Hutwagner et al. 2004;

De Jong, Andersson et al. 2005; Finley, Reid-Smith et al. 2006) and human to human transmission has also been described (Todd, Greig et al. 2008). Those sources are not monitored actively, but as for non-animal food sources, some data are available through the *Salmonella* Network, which definitely offers a useful alternative to active surveillance for the sources, whatever they are. Finally, some private initiatives exist such as Qualimat (<http://www.qualimat.org/>), an association which runs an auto-control plan for *Salmonella* in raw materials with the participation of the feed manufacturers, and as the PFGE profiles database managed by the “Institut du Porc” (http://www.ifip.asso.fr/actu/pdf/MA_aval.pdf). This shows a global concern about *Salmonella* and a will to monitor and to control it covering all sectors, imperfectly, but efficiently. Because of these reasons, the multiplicity of the systems and actors can be revealed to be a positive point.

The success of integrated surveillance at a national level depends on the availability of high-quality, accessible and comparable data used for meaningful analysis and reporting. Such an effective surveillance system should improve the scientific basis, on the one hand, for the implementation of management measures and, on the other hand, for the evaluation of their impact and risk analysis.

Thus, the issues to be addressed by a surveillance system are diverse: detection of outbreaks, burden of illness and economic burden estimation, prevention, assessment of the impact of interventions, microbiological risk assessment (Desenclos, Vaillant et al. 2002). Close collaboration between the different actors in surveillance, public health and food safety authorities and the harmonization of methodologies between the human and agro-food sectors is a key point in addressing most of these issues (Batz, Doyle et al. 2005; Ammon and Tauxe 2007; Sofos 2008) and seems to be a success in France.

If passive networks often suffer from reporting delay and selection biases, their daily computerized databases are suited to an alert activity. The NRC performs weekly unusual event detection by time series analysis on its database. This allows the public health authorities concerned to be informed quickly and to trigger the investigation of the outbreaks detected. The epidemiological investigations coordinated by InVS rely on combined data across agro-food and human surveillance (namely the NRC and the *Salmonella* Network), which is made possible thanks to the harmonization of the laboratory methodologies. A similar approach is being implemented (Baroukh, Le Strat et al. 2008) by

the *Salmonella* Network which could allow earlier warnings, from the contamination of the food-chain onwards and thus possibly prevent outbreaks.

Sporadic cases require a different approach, based on prevention and risk-based interventions. Attributing the cases to the main food-animal reservoirs is of the utmost importance in conceiving and assessing the impact of such interventions ((ICMSF) 2006; Ammon and Tauxe 2007; Sofos 2008). Several approaches can be used to perform attribution. In the case of sporadic case attribution, it is necessary to dispose of representative data in the agro-food sector. Currently, a microbial subtyping attribution project is in progress in France (David, Danan et al. 2008). This approach could be implemented thanks to the active collaboration of all the different surveillance actors.

Based on the information provided by the different surveillance systems, management actions are implemented and should be evaluated regularly according to outcome indicators. In France, interventions have been enforced at farm level. The first line impact is the level of prevalence in flocks, but the real objective is the number of cases induced. Within the framework of European regulation, continuous control programs are implemented with prevalence mitigation targets defined according to the animal production concerned. In France, the impact of one-year prophylaxis in *Gallus gallus* demonstrated a reduction of the prevalence of *Salmonella* in breeding flocks from 3% in 2006 to 0.6% in 2007 for the five regulated serotypes. These results show the efficiency of the programs and the good level of sanitary conditions of breeding production in *Gallus gallus*. Furthermore, the impact of the national control program implemented in the breeding flocks of *Gallus gallus* from 1998 onwards on human cases has been evaluated through time series analysis by Poirier et al (Poirier, Watier et al. 2008), using NRC data. A 33% decrease was observed in the prevalence in the flocks, and a significant decrease in human cases was demonstrated between 1998 and 2003, up to 21% for cases associated with the serotype Enteritidis.

Finally, collaboration also operates on an international level. The French human public health and food safety authorities cooperate at European level with ECDC and EFSA, and the NRC is WHO collaborating centre for reference and research on *Salmonella*.

References

- Anonymous, 2006: Report of Task Force on Zoonoses Data Collection on Guidance Document on Good Practices for Design of Field Surveys. *The EFSA Journal*, **93**, 1-24.
- Anonymous, 2009: The community Summary Report on Trends and Sources of Zoonoses and Zoonotic Agents in the European Union in 2007. *The EFSA Journal*, **223**.
- Allard, D. G., 2002: The 'farm to plate' approach to food safety - Everyone's business. *Canadian Journal of Infectious Diseases*, **13**, 185-190.
- Ammon, A. & R. V. Tauxe, 2007: Investigation of multi-national foodborne outbreaks in Europe: some challenges remain. *Epidemiol Infect*, **135**, 887-889.
- Angulo, F. J., A. C. Voetsch, D. Vugia, J. L. Hadler, M. Farley, C. Hedberg, P. Cieslak, D. Morse, D. Dwyer & D. L. Swerdlow, 1998: Determining the burden of human illness from food borne diseases. CDC's emerging infectious disease program Food Borne Diseases Active Surveillance Network (FoodNet). *Vet Clin North Am Food Anim Pract*, **14**, 165-172.
- Baroukh, T., Y. Le Strat, F. Moury, A. Brisabois & C. Danan, 2008: Use of statistical methods for a routinely unusual event surveillance in non human Salmonella. *European Scientific Conference on Applied Infectious Disease Epidemiology (ESCAIDE)*. Berlin.
- Batz, M. B., M. P. Doyle, J. G. Morris Jr., J. Painter, R. Singh, R. V. Tauxe, M. R. Taylor & D. M. A. L. F. Wong, 2005: Attributing illness to food. *Emerg. Infect. Dis.*, **11**, 993-999.
- Botrel, M. A., M. Chazel, D. Meunier, E. Jouy, M. Kobisch, J. Y. Madec & D. Calavas, 2006: Le RESAPATH : Analyse critique et propositions d'amélioration. *Epidémiologie et santé animale*, **50**, 157-168.
- Brandl, M. T., 2006: Fitness of human enteric pathogens on plants and implications for food safety. *Annu Rev Phytopathol*, **44**, 367-392.
- Brouard, C., E. Espié, F. X. Weill, A. Kerouanton, A. Brisabois, A. M. Forgue, V. Vaillant & H. de Valk, 2007: Two consecutive large outbreaks of Salmonella enterica serotype Agona infections in infants linked to the consumption of powdered infant formula. *Pediatr Infect Dis J*, **26**, 148-152.
- CDC, 2001: Updated guidelines for evaluating public health surveillance systems. *MMWR*, **50(RR13)**, 1-35.
- Chazel, M., Y. Buret, D. Meunier, J. Y. Madec & D. Calavas, 2007: Le RESSAB - Réseau d'EpidémiologieSurveillance des Salmonelloses Bovines - Résultats 2006. *Bulletin Epidémiologique - AFSSA*, **25**, 5-7.
- David, J., C. Danan, F. X. Weill, P. Fravallo, N. Bemrah-Aouachria, J. Y. Madec, M. Picherot, N. Jourdan Da Silva, A. Brisabois, P. Sanders & D. Guillemot, 2008: Developing a model to quantify the contribution of the major food-animal sources to human cases of salmonellosis regarding their sensibility to antimicrobials in France. *4th Med-Vet-Net annual scientific meeting*. Saint Malo, France.

- De Jong, B., Y. Andersson & K. Ekdahl, 2005: Effect of regulation and education on reptile-associated salmonellosis. *Emerg. Infect. Dis.*, **11**, 398-403.
- Desenclos, J. C., V. Vaillant & H. De Valk, 2002: [Foodborne infections: do we need further data collection for public health action and research]. *Rev Epidemiol Sante Publique*, **50**, 67-79.
- Dominguez, M., N. Jourdan-Da Silva, V. Vaillant, N. Pihier, C. Kermin, F. X. Weill, G. Delmas, A. Kerouanton, A. Brisabois & H. de Valk, 2008: Outbreak of Salmonella enterica Serotype Montevideo Infections in France Linked to Consumption of Cheese Made from Raw Milk. *Foodborne Pathog Dis*.
- Dufour, B. & S. La Vieille, 2000: Epidemiological surveillance of infectious diseases in France. *Vet Res*, **31**, 169-185.
- Ellerbroek, L., T. Alter, R. Johne, K. Nockler, L. Beutin & R. Helmuth, 2009: [Tasks and duties of veterinary reference laboratories for food borne zoonoses]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, **52**, 208-213.
- Espié, E., H. De Valk, V. Vaillant, N. Quelquejeu, F. Le Querrec & F. X. Weill, 2005b: An outbreak of multidrug-resistant Salmonella enterica serotype Newport infections linked to the consumption of imported horse meat in France. *Epidemiol Infect*, **133**, 373-376.
- Espié, E., F. X. Weill, C. Brouard, I. Capek, G. Delmas, A. M. Forgues, F. Grimont & H. de Valk, 2005a: Nationwide outbreak of Salmonella enterica serotype Agona infections in infants in France, linked to infant milk formula, investigations ongoing. *Euro Surveill*, **10**, E050310 050311.
- Finley, R., R. Reid-Smith & J. S. Weese, 2006: Human health implications of Salmonella-contaminated natural pet treats and raw pet food. *Clin Infect Dis*, **42**, 686-691.
- Flint, J. A., Y. T. Van Duynhoven, F. J. Angulo, S. M. DeLong, P. Braun, M. Kirk, E. Scallan, M. Fitzgerald, G. K. Adak, P. Sockett, A. Ellis, G. Hall, N. Gargouri, H. Walke & P. Braam, 2005: Estimating the burden of acute gastroenteritis, foodborne disease, and pathogens commonly transmitted by food: An international review. *Clin. Infect. Dis.*, **41**, 698-704.
- Gallay, A., V. Vaillant, P. Bouvet, P. Grimont & J.-C. Desenclos, 2000: How many foodborne outbreaks of Salmonella infection occurred in France in 1995? *Am. J. Epidemiol.*, **152**, 171-177.
- Gerner-Smidt, P., K. Hise, J. Kincaid, S. Hunter, S. Rolando, E. Hyttia?-Trees, E. M. Ribot, B. Swaminathan, M. Adams, M. Bird, P. Bordoni, K. Cooper, S. Van Duyn, M. A. Fair, N. Garrett, L. M. Graves, D. Jennings, M. Joyner, B. McGlinchey & N. Patel, 2006: PulseNet USA: A five-year update. *Foodborne Pathog. Dis.*, **3**, 9-19.
- Gervelmeyer, A., M. Hempen, U. Nebel, C. Weber, S. Bronzwaer, A. Ammon & P. Makela, 2008: Developing the Community reporting system for foodborne outbreaks. *Euro Surveill*, **13**, pii: 19029.
- Haeghebaert, S., P. Sulem, L. Deroudille, E. Vanneroy-Adenot, O. Bagnis, P. Bouvet, F. Grimont, A. Brisabois, F. Le Querrec, C. Hervy, E. Espie, H. de Valk & V. Vaillant, 2003: Two outbreaks of Salmonella enteritidis phage type 8 linked to the consumption of Cantal cheese made with raw milk, France, 2001. *Euro Surveill*, **8**, 151-156.
- Hald, T., D. Vose, H. C. Wegener & T. Koupeev, 2004: A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Anal*, **24**, 255-269.
- Hendriks, P., M. Chazel, Y. Buret & B. Dufour, 2005: Development of performance indicators for the bovine clinical salmonellosis surveillance network in France. *J Vet Med B Infect Dis Vet Public Health*, **52**, 465-475.

- Herida, M., 2008: Evaluation du dispositif des maladies à déclaration obligatoire en France : connaissances, attitudes et pratiques des médecins et des biologistes, 2005. *Bulletin épidémiologique hebdomadaire*, **51-52**, 503-507.
- Hueston, W. D., 1993: Assessment of national systems for the surveillance and monitoring of animal health. *Rev Sci Tech*, **12**, 1187-1196.
- ICMSF., 2006: Use of epidemiologic data to measure the impact of food safety control programs. *Food Control*, **17**, 825-837.
- InVS, 2004: Morbidité et mortalité dues aux maladies infectieuses d'origine alimentaire en France. InVS.
- Jourdan, N., S. Le Hello, G. Delmas, J. Clouzeau, C. Manteau, B. Desaubliaux, V. Chagnon, F. Thierry-Bled, N. Demare, F. Weill & H. De Valk, 2008: Nationwide outbreak of Salmonella enterica serotype Give infections in infants in France, linked to infant milk formula, September 2008. *Euro Surveill*, **13**.
- Jouy, E., D. Meunier, M. Chazel, I. Kempf, M. Kobisch, J. P. Orand & P. Sanders, 2003: RESAPATH : Réseau de surveillance de la résistance aux antibiotiques chez les principales bactéries pathogènes des bovins, des porcs et des volailles. Description des espèces bactériennes et des pathologies enregistrées en 2002. *Bulletin Epidémiologique - AFSSA*, **9**, 5-6.
- Kerouanton, A., M. Marault, R. Lailier, F. X. Weill, C. Feurer, E. Espie & A. Brisabois, 2007: Pulsed-field gel electrophoresis subtyping database for foodborne Salmonella enterica serotype discrimination. *Foodborne Pathog Dis*, **4**, 293-303.
- Lailier, R., M. Sanaa, J. Chadoeuf, B. Fontez, A. Brisabois, C. Colmin & Y. Millemann, 2005: Prevalence of multidrug resistant (MDR) Salmonella in bovine dairy herds in western France. *Prev. Vet. Med.*, **70**, 177-189.
- Mead, P. S., L. Slutsker, V. Dietz, L. F. McCaig, J. S. Bresee, C. Shapiro, P. M. Griffin & R. V. Tauxe, 1999: Food-related illness and death in the United States. *Emerg. Infect. Dis.*, **5**, 607-625.
- Mermin, J., L. Hutwagner, D. Vugia, S. Shallow, P. Daily, J. Bender, J. Koehler, R. Marcus & F. J. Angulo, 2004: Reptiles, amphibians, and human Salmonella infection: A population-based, case-control study. *Clin. Infect. Dis.*, **38**.
- Molbak, K., 2005: Human health consequences of antimicrobial drug-resistant Salmonella and other foodborne pathogens. *Clin Infect Dis*, **41**, 1613-1620.
- Moury, F., S. Fremy & A. Brisabois, 2006: Épidémiosurveillance des salmonelles d'origine non humaine : données du réseau salmonella - année 2005. *Bulletin épidémiologique - AFSSA*, **20**, 6-8.
- Poirier, E., L. Watier, E. Espie, F. X. Weill, H. De Valk & J. C. Desenclos, 2008: Evaluation of the impact on human salmonellosis of control measures targeted to Salmonella Enteritidis and Typhimurium in poultry breeding using time-series analysis and intervention models in France. *Epidemiol Infect*, **136**, 1217-1224.
- Singer, R. S., L. A. Cox Jr., J. S. Dickson, H. S. Hurd, I. Phillips & G. Y. Miller, 2007: Modeling the relationship between food animal health and human foodborne illness. *Prev. Vet. Med.*, **79**, 186-203.
- Sofos, J. N., 2008: Challenges to meat safety in the 21st century. *Meat Sci.*, **78**, 3-13.
- Souillard, R., J. Toux, S. Le Bouquin & V. Michel, 2007: Le RNOEA : Réseau National d'Observations Epidémiologiques en Aviculture Pathologie aviaire en 2004. *Bulletin Epidémiologique - AFSSA*, **24**, 6-7.

- Stark, K. D., G. Regula, J. Hernandez, L. Knopf, K. Fuchs, R. S. Morris & P. Davies, 2006: Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. *BMC Health Serv Res*, **6**, 20.
- Todd, E. C., J. D. Greig, C. A. Bartleson & B. S. Michaels, 2008: Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 5. Sources of contamination and pathogen excretion from infected persons. *J Food Prot*, **71**, 2582-2595.
- Vaillant, V. & E. Espie, 2005: Cases of Salmonella Stourbridge infection in France, April-June 2005. *Euro Surveill*, **10**, E050721 050722.
- Weill, F.-X., F. Guesnier, V. Guibert, M. Timinouni, M. Demartin, L. Polomack & P. A. D. Grimont, 2006: Multidrug resistance in Salmonella enterica serotype typhimurium from humans in France (1993 to 2003). *J. CLIN. MICROBIOL.*, **44**, 700-708.
- Woodward, D. L., R. Khakhria & W. M. Johnson, 1997: Human salmonellosis associated with exotic pets. *J. CLIN. MICROBIOL.*, **35**, 2786-2790.

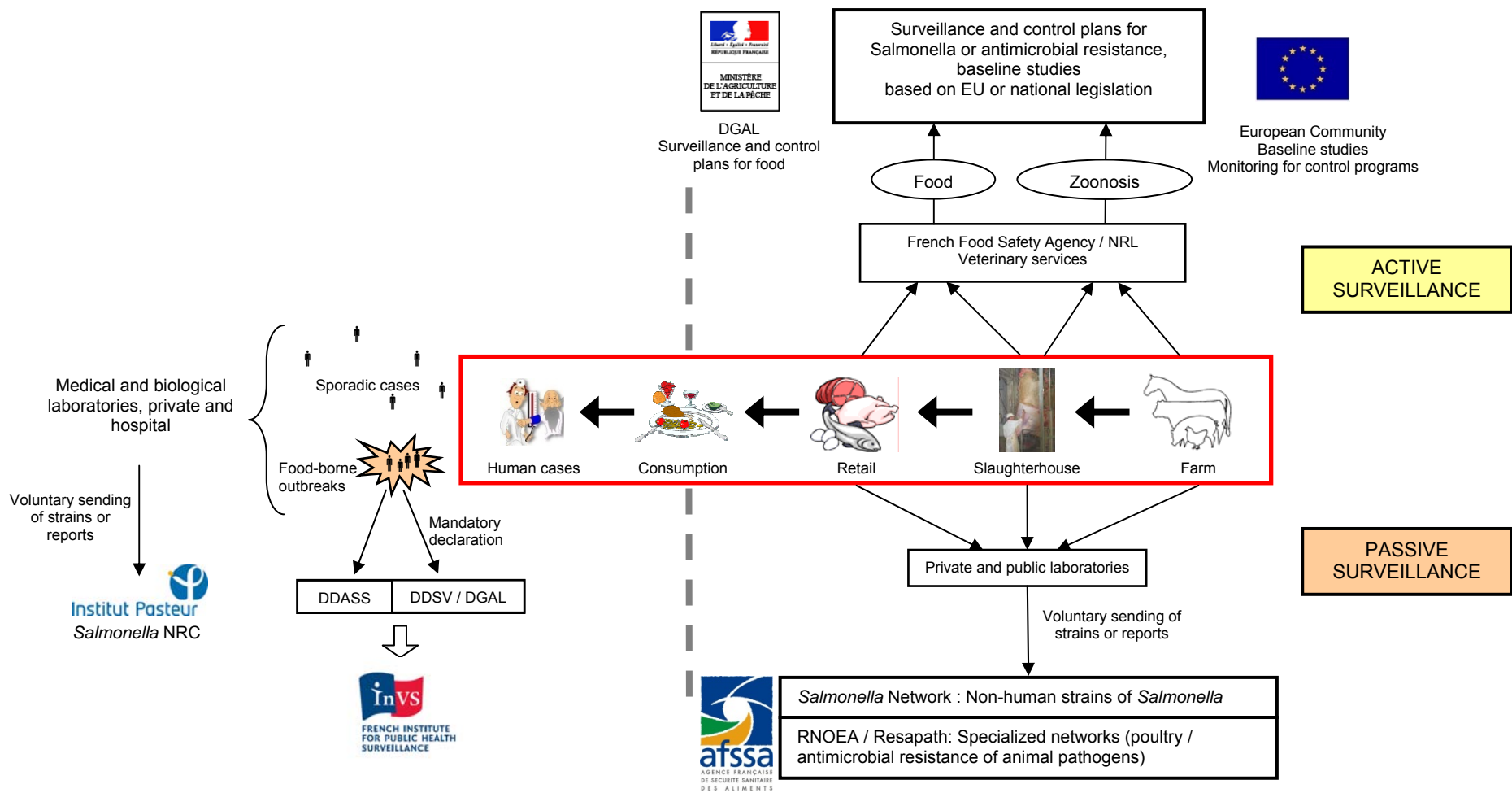


Figure 1: Organization of the *Salmonella* surveillance system in France in 2009

	National public health network: foodborne outbreak	National Reference Centre for Salmonella
Coordinator	InVS	NRC
Objectives	Registration of foodborne outbreak events with the biological agent, the food item involved and the number of cases	Temporal trends in serotypes and AMR of human cases Alerts in case of unusual events (number of cases)
Means of collection	Passive surveillance	Passive surveillance
Scope of surveillance	National	National
Monitoring period	Continuous since 1987	Continuous since 1947
Pivotal variable ^a	Number of outbreaks, number of cases per outbreak	Number of cases per serotype
Disease focus	General - multi-pathogens	<i>Salmonella</i> specific
Design	Descriptive + risk factors	Descriptive
Temporality	Trends	Trends
Availability of non pivotal information	Yes	Yes
Way of selection	Non-random	Non-random
Source of data	Mandatory disease notification	Laboratory investigation records and biological specimen banks based on an opportunistic collection
Coverage of the population	OK (mandatory).	2006: 1357 laboratories, among which 1028 private laboratories and 329 hospital laboratories, 30 to 40% of French clinical laboratories
Duplicates	Eliminated	Eliminated
Target population	Foodborne outbreaks occurring in France	<i>Salmonella</i> infected people
Study population	Foodborne outbreaks occurring in France that are notified	Cases confirmed by a laboratory collaborating with the NRC
Epidemiological unit	Outbreak	Strain
Case definition	Illness in at least two persons with digestive symptoms that can be attributed to the same food source	<i>Salmonella</i> strain identified in the sample
Laboratory testing	Not mandatory. For <i>Salmonella</i> confirmed outbreaks: private or hospital laboratory for identification of human cases of salmonellosis, NRC for serotyping. For veterinary isolates: LVD for identification and AFSSA for serotyping	Identification / Serotyping / Resistance phenotype (Disk Diffusion method, E-test) / Genotype (PCR, PFGE, MLVA, BLSE)
Data collection and management	DDASS, DDSV, InVS	NRC
Analytical methodologies	Descriptive. Case-control studies or cohort studies can be conducted using data that do not exist in the mandatory notification of foodborne outbreak	Descriptive Specific for unusual event detection
Quality assurance	For missing or invalid data, further enquiries at the DDSV or the DDASS	Internet quality control "Quality approach" of the Pasteur Institute
Bias	Importance of the clinical symptoms Availability of samples to identify the biological agent Diffuse outbreaks Family outbreaks are less reported than collective or commercial catering outbreaks	Importance of the clinical symptoms Willingness of the laboratories to participate Population groups at risk
Reporting	Report: Weekly epidemiological bulletin, scientific publications	Annual report, scientific publications

Table 1: Surveillance of human cases in 2008

a: variable describing the quantity corresponding to the main objective of the survey.

	European control programs	European baseline studies
Coordinator	DGAL	DGAL
Objectives	Control of <i>Salmonella</i> prevalence in poultry	Evaluation of the prevalence level of <i>Salmonella</i> at farm level or at the slaughterhouse level
Means of collection	Active surveillance	Active surveillance
Scope of surveillance	National	National
Monitoring period	Continuous since 2007 (1998 for Enteritidis (SE) and Typhimurium (ST) in breeding flocks)	12 months, from 2004 to 2009 according to the channel
Pivotal variable	Prevalence: SE, ST, Hadar, infantis, Virchow for breeding flocks (laying hens and broilers) and SE, ST for production flocks of laying hens and broilers	Prevalence per serotype
Disease focus	<i>Salmonella</i> specific	<i>Salmonella</i> specific
Design	Descriptive + risk factors	Descriptive + risk factors
Temporality	Trends	Single occurrence
Availability of non pivotal information	Yes	Yes
Way of selection	Structured	Structured
Methods of selection	Exhaustive	Random selection Stratification on the farm or slaughterhouse size
Representativeness	Not relevant	OK
Target population	Breeding flocks (from 2007 on), laying hens (from 2008 on), broilers (from 2009 on)	Production and breeding flocks of laying hens, broilers, pigs and turkeys Flocks or animals either at farm level and/or at the slaughterhouse
Study population	Registered farms Breeding flocks: over 250 animals Production flocks of laying hens: over 250 animals delivered to a conditioning unit Production flocks of broilers: all the flocks except under 250 directly delivering to the consumer	Laying hens, broilers, pigs and turkeys farms or slaughterhouses Farms registered with a minimal size Slaughterhouses with a minimal size
Epidemiological unit	Building	Flock or herd
Case definition	At least 1 positive sample for broilers, at least 2 consecutive positive samples for breeders and laying hens	At least 1 positive sample
Laboratory testing	Laboratories with accreditation isolation / identification / serotyping (NFU 47100)	NRL for <i>Salmonella</i> Isolation / identification / serotyping
Data collection and management	DGAL / Veterinary Services Direction	DGAL / NRL
Analytical methodologies	Descriptive (investigation if contamination)	EFSA, Specific / NRL at national level
Quality assurance	Certified laboratories	NRL / certified laboratories
Validation (bias)	Small flocks not included Direct selling of broilers not included Multicentric (investigator bias)	Small flocks and slaughterhouses not included Detection (a few units per building or band at the slaughterhouse)
Reporting	Annual Community report Intern communication	EFSA reports / NRL scientific publications

Table 2: Surveillance of food-animals in 2008

	Surveillance plans	Control plans
Coordinator	DGAL	DGAL
Objectives	Evaluation of the prevalence level of <i>Salmonella</i> in carcasses and products Specific questions such as comparison of sampling methods	Evaluation of the prevalence level of <i>Salmonella</i> on carcasses and products
Means of collection	Active surveillance	Active surveillance
Scope of surveillance	National	National
Monitoring period	several months to one year	several months to one year for broiler meat
Pivotal variable	prevalence (spp or per serotype)	prevalence per serotype
Disease focus	<i>Salmonella</i> specific	<i>Salmonella</i> specific
Design	Descriptive	descriptive
Temporality	Single occurrence	single occurrence
Availability of non pivotal information	No	No
Way of selection	Structured	Structured
Methods of selection	Random selection of slaughterhouses stratified on the size Random selection of the units	Random selection of the slaughterhouses or producers or targeted selection according to specific activities
Representativeness	OK	OK
Target population	Carcasses from the targeted channels Products in approved firms	Carcasses from the targeted channels Meat products: poultry, minced and mechanically separated meat
Study population	Food animals slaughtered in the selected slaughterhouses, products or retail units available in the targeted firms (producers or retail)	Pigs, large cattle and sheep slaughtered in the selected slaughterhouses Products in the selected factories
Epidemiological unit	Carcass, retail unit	Carcass, batch
Case definition	At least 1 positive sample	At least 1 positive sample
Laboratory testing	Laboratories with accreditation or LDA and AFSSA isolation / identification / serotyping (ISO 6579)	LVD isolation / identification / serotyping
Data collection and management	DGAL	DGAL
Analytical methodologies	Descriptive and specific	Descriptive
Quality assurance	COFRAC accreditation for meat products ISO 6579	COFRAC accreditation for meat products ISO 6579
Validation (bias)	Large flocks at higher probability of being sampled Multicentric (investigator bias)	Large flocks at higher probability of being sampled Multicentric (investigator bias)
Reporting	Memos, internal communication Report (synthesis) on the agriculture ministry website	Memos, internal communication Report (synthesis) on the agriculture ministry website

Table 3: Surveillance of food in 2008

	<i>Salmonella</i> Network	Resapath	RNOEA
Coordinator	AFSSA	AFSSA	AFSSA
Objectives	Spatio-temporal trend analyses of <i>Salmonella</i> throughout the food chain and detection of unusual events	Spatio-temporal trend analyses and emergence detection of antimicrobial resistance in veterinary pathogenic bacteria	Spatio-temporal trend analyses and emergence detection of poultry diseases
Means of collection	Passive surveillance	Passive surveillance	Passive surveillance
Scope of surveillance	National	National	National
Monitoring period	Continuous, since 1997	Continuous, since 1982	Continuous, since 1987
Pivotal variable	Strain number per serotype	Strain number per animal channel and pathology	Notification number (contaminated flocks and/or disease diagnostic notifications)
Disease focus	<i>Salmonella</i> specific	Multi-pathogens	Multi-pathogens
Design	Descriptive	Descriptive (+ transversal studies)	Descriptive
Temporality	Trends	Trends	Trends
Availability of non pivotal information	Yes	Yes	Yes
Way of selection	Non-random	Non-random	Non-random
Source of data	Laboratory investigation records and biological specimen banks based on an opportunistic collection	Laboratory investigation records based on an opportunistic collection	Vet and laboratory investigation records based on an opportunistic collection
Coverage of the population	97% and 77% respectively of public and private veterinary laboratories	59 laboratories in 52 departements	60 vets and laboratories
Duplication of data	Not controlled	Not controlled	Controlled
Target population	All animal productions, foodstuffs and environment	All animals sampled for an antibiogram	National poultry production
Study population	Strains from animals, foodstuffs and environment analysed by participating laboratories	Animals sampled for an antibiogram and analysed by the network partner laboratories (food animals and pets)	Poultry flocks traced by vets or laboratories
Epidemiological unit	Strain	Strain	Flock
Case definition	<i>Salmonella</i> strain identified in the sample	Strain isolated from an ill animal	Flock affected by a disease and notified by vets
Laboratory testing	Identification / Serotyping: laboratories and AFSSA AMR and PFGE: AFSSA	Identification / serotyping / AMR: participating laboratories	Identification / serotyping: participating laboratories
Data collection and management	AFSSA – LERQAP / CEB	AFSSA Lyon and Ploufragan	AFSSA Ploufragan
Analytical methodologies	Descriptive statistical analyses, unusual event detection	Descriptive statistical analyses + specific	Descriptive statistical analyses
Quality assurance	Quality assurance of LERQAP Accreditation for serotyping, Inter-laboratory trial for serotyping	Quality assurance of AFSSA laboratories, inter-laboratory trial for AMR testing	Standardization of poultry diseases diagnosis by common guidelines No analytical standardization
Bias	Willingness of the laboratories to participate Impact of regulation Investigator bias	Willingness of the laboratories to participate Detection bias Investigator bias	Willingness of the laboratories to participate Declaration bias Investigator bias Impact of the regulation
Reporting	Quarterly and annual reports to the network partners, unusual event reporting	Annual reports to the network partners	Bi-monthly and annual reports to the network partners

Table 4: Non-human passive surveillance in 2008

II

Sélection des données et adaptation à l'utilisation en attribution

La démarche d'attribution par typage microbiologique nécessite de disposer, pour une année donnée, de trois types d'informations : le nombre de cas de salmonelloses humaines, le niveau de contamination par les salmonelles des sources animales considérées et la consommation de ces dernières en France (Hald, Vose et al. 2004).

La première source d'information doit permettre d'identifier les nombres de cas domestiques sporadiques (i.e. cas ayant été infectés sur le territoire national et qui ne sont pas liés à un foyer épidémique) par type microbiologique. Il est donc nécessaire de savoir pour chacun des cas, en plus du type microbiologique de la salmonelle à l'origine de l'infection, si celui-ci est lié à un foyer de cas groupés et si la salmonelle a pu être acquise à l'étranger. Pour identifier le niveau de consommation, l'information doit couvrir l'ensemble des sources sélectionnées et être disponible pour la même période de temps que les souches recueillies en surveillance humaine et en surveillance animale. Dans ce travail, la consommation considérée est celle de l'ensemble de la population exposée, ici, la population française. La contamination animale est mesurée sous forme de prévalence ou à défaut, de répartition des types microbiologiques au sein de chacune des sources considérées. Le point de mesure de la contamination des sources par les salmonelles se situe le plus en amont possible au sein des filières, à savoir en élevage et/ou sur carcasses en abattoir. En effet travailler en amont permet d'éviter, en particulier, les problématiques de contamination croisée.

Nous avons identifié parmi les acteurs de la surveillance française présentés au chapitre précédent, ceux qui détenaient les données et informations nécessaires pour appliquer un modèle d'attribution tel que celui par typage microbiologique. Le CNR et l'InVS ont été sollicités pour la partie humaine, la DGAL, qui coordonne les plans de surveillance européens et nationaux, l'a été pour les données animales ainsi que le Réseau Salmonella qui fournit des données de surveillance passive. Enfin les données de consommation proviennent de l'AFSSA, de l'Office national Interprofessionnel des produits de la mer et de l'aquaculture (OFIMER) et de l'Office de l'Elevage (récemment regroupés avec d'autres offices au sein de FranceAgriMer). Les données considérées n'ont pas été collectées dans l'objectif d'une utilisation en attribution et ne correspondaient donc pas toujours aux exigences fixées par les concepteurs de l'outil d'attribution. Un travail d'adaptation de ces données a donc été réalisé.

Les paragraphes suivants présentent donc pour chaque type de données (humaines, animales, consommation), l'identification et l'évaluation des données de surveillance existantes, puis le travail effectué pour aboutir à un jeu de données exploitable dans une démarche d'attribution. Le mode de sous-typage et la sélection des sérotypes pour le modèle seront également abordés. En raison de la disponibilité des données animales et humaines, l'année de référence considérée pour ces travaux est 2005.

1 Données humaines

On cherche à déterminer les nombres de cas de salmonelloses non typhiques domestiques et sporadiques par type microbiologique pour l'année 2005.

1.1 Données disponibles

Il n'existe pas d'obligation en France de déclaration des cas de salmonellose non groupés, aussi un inventaire exhaustif des cas humains n'est-il pas disponible.

Sources de données

Les données utilisées ont été fournies par le CNR et concernent des cas confirmés de salmonelloses. Des données complémentaires concernant les cas groupés, nécessaires pour écarter les cas non sporadiques de la base de données, ont été fournies par l'InVS, à partir des déclarations de Toxi-Infections Alimentaires Collectives (TIAC).

Informations disponibles¹

Le CNR centralise soit des souches qui sont analysées (il y sera fait référence sous le terme de souchier), soit des récapitulatifs d'analyse concernant des souches isolées au sein des laboratoires. Lors de l'envoi de récapitulatifs, seul le sérotype est renseigné. Pour les cas issus du souchier, en sus du sérotype, les informations suivantes sont renseignées : les dates d'isolement et de réception, l'enregistrement d'un voyage récent, les données générales

¹ Seules les informations utilisées pour l'adaptation des données sont indiquées

concernant le patient (âge, localisation géographique) et localisation géographique du laboratoire d'où provient la souche. Concernant les TIAC enregistrées à l'InVS, nous avons utilisé les informations concernant le nombre de cas, la localisation géographique, les dates de l'épisode ainsi que l'agent suspecté ou confirmé et son sérotype (si disponible) et enfin, le numéro CNR des cas liés à une TIAC investiguée.

Ainsi, pour l'année 2005 ont été recensés : 11 212 cas humains dont 6 551 issus du souchier et 4 661 issus de récapitulatifs répartis en 268 sérotypes.

Pour la même année, 174 TIAC liées à *Salmonella* ont été enregistrées à l'InVS. Sur ces 174 TIAC, 33% sont de sérotype inconnu, 40% sont de sérotype Enteritidis et 24% sont de sérotype Typhimurium. Enfin, 5 épidémies diffuses nationales ont été investiguées par l'InVS. Les souches identifiées étaient Agona, Worthington, Manhattan, Stourbridge et Oranienburg.

1.2 Adaptation des données pour l'attribution

Ce travail porte tout d'abord sur les cas issus du souchier. Les cas typhiques sont d'abord écartés, soit 279 cas de sérotype Typhi ou Paratyphi. Pour obtenir des cas sporadiques, ont été de plus exclus : 21 doublons « complets »¹ (souche enregistrée plusieurs fois), 241 échantillonnages multiples¹ (même patient, même sérotype, dates d'échantillonnage différentes), 147 cas liés à des foyers familiaux¹ (même nom de famille, même sérotype, coïncidence spatio-temporelle) et 160 cas liés à des TIAC. Ces derniers comprennent les cas de la base de données enregistrés à l'InVS comme liés aux 5 épidémies investiguées, ainsi que des cas retrouvés par coïncidence spatio-temporelle (les données géographiques concernant le patient ou le laboratoire) avec les TIAC de sérotype autre que Typhimurium et Enteritidis. En effet, ces deux sérotypes sont très fréquents ce qui rend la détection de coïncidence spatio-temporelle difficile. Un cas est conservé dans la base de données pour représenter chaque groupe de doublons, chaque foyer familial et chaque TIAC. Enfin, pour obtenir des cas domestiques, 336 cas pour lesquels un voyage était renseigné et 358 cas enregistrés dans les Départements et Territoires d'outre-mer (Dom Tom) ont été exclus. Au total, et après exclusion de 20% des cas non typhiques, 5 035 cas non typhiques estimés domestiques sporadiques ont été identifiés dans le souchier (un cas pouvant être concerné par plusieurs motifs d'exclusion) (Figure 2.1).

¹ Ces opérations ont été effectuées au sein du CNR, sur la base de données avant anonymisation, par recouplement des données géographiques, temporelles et nominatives.

Concernant les récapitulatifs, 4 661 cas non typhiques sont recensés. Sous l'hypothèse que les cas issus du souchier et des récapitulatifs ont les mêmes caractéristiques, les pourcentages d'exclusion estimés à partir du souchier ont été appliqués aux récapitulatifs en stratifiant sur le sérotype. Un total de 620 cas a été exclu, soit 13% des cas non typhiques. Ainsi, 4 041 cas non typhiques domestiques sporadiques ont été identifiés pour les cas issus des récapitulatifs.

1.3 Résultats

In fine, 9 076 cas répartis entre 230 sérotypes différents ont été retenus. Les sérotypes Typhimurium et Enteritidis, représentent 74% des cas (39% et 35% respectivement), le troisième sérotype par ordre d'importance, Infantis, représente seulement 2% des cas (Cf figure 2.2). Deux sérotypes sont donc très majoritaires, les autres sérotypes sont de moindre importance, mais nombreux (266 sérotypes pris en compte).

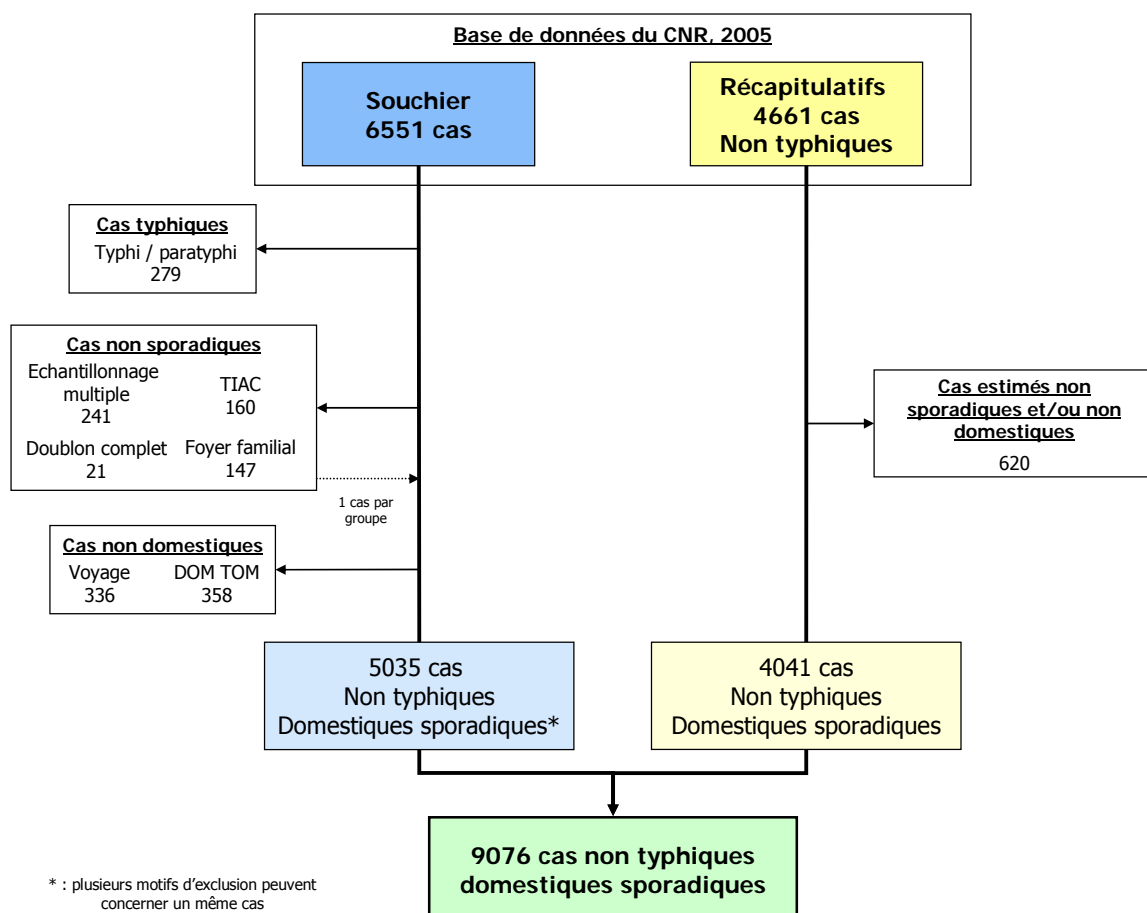


Figure 2.1 : Résumé des critères d'exclusion et effectifs
 Pour des raisons de lisibilité, les sérotypes de très faible représentativité ($n < 10$) sont rassemblés sur le graphe dans une catégorie « autres ».

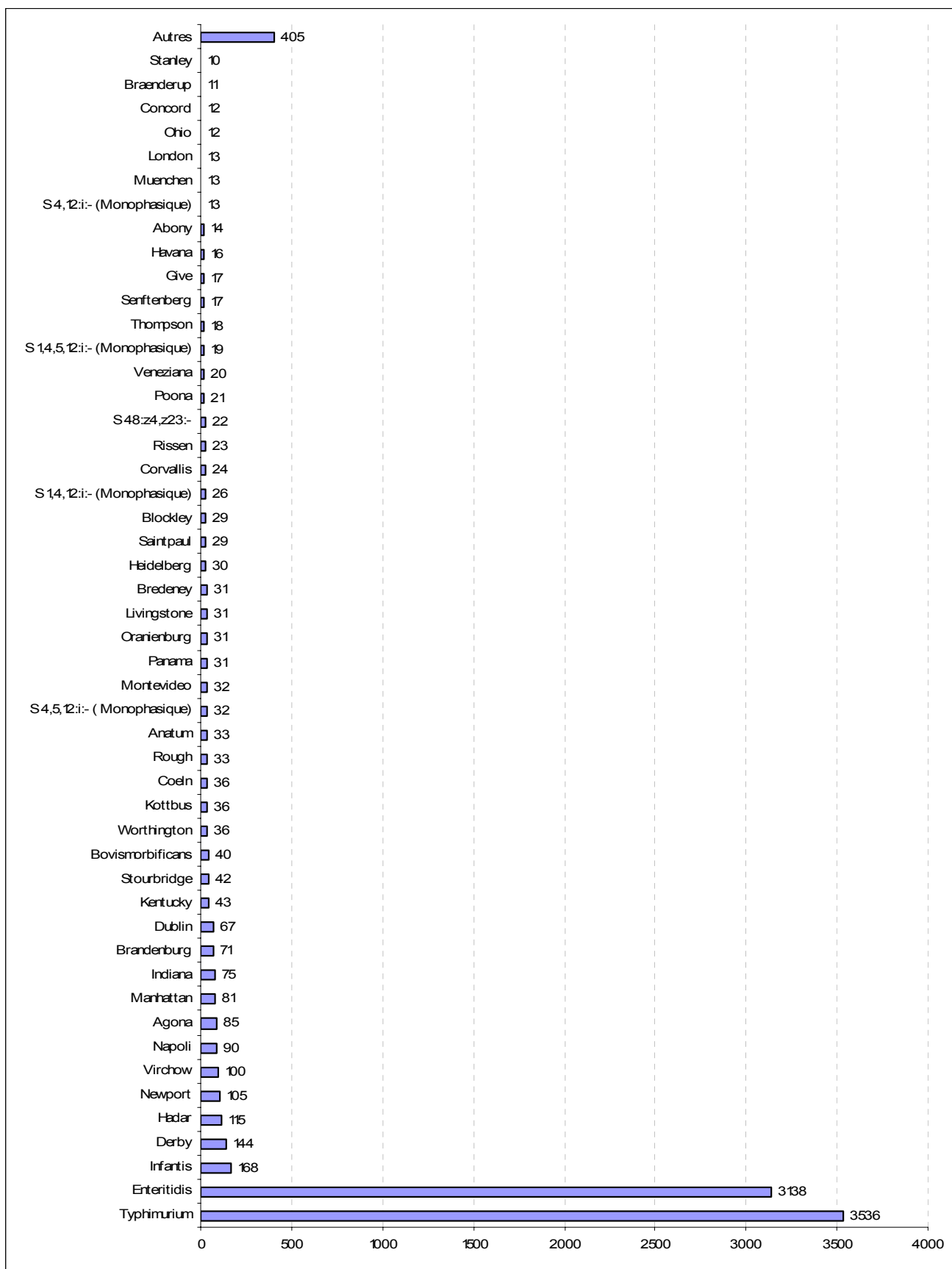


Figure 2.2 : Répartition par sérotype des cas humains domestiques sporadiques non typhiques (n=9 076, données CNR de 2005)

2 Données animales

Le modèle nécessite des données de prévalence ou, à défaut, de répartition des différents sérotypes dans les filières en 2005. La démarche d'attribution par typage microbiologique nécessite pour les sources de recueillir les données le plus en amont possible, c'est-à-dire en élevage ou à l'abattoir.

2.1 Données disponibles

En France, deux systèmes de surveillance des salmonelles dans les filières animales existent en parallèle : la surveillance active institutionnelle sur base réglementaire européenne ou nationale et une surveillance passive basée sur la centralisation de souches non humaines issues de laboratoires d'analyses vétérinaires privés ou publics volontaires (David, Danan et al. submitted).

2.1.1 Données de surveillance active

Sources de données

Les données de surveillance active retenues dans ce travail ont été mises à disposition par la DGAL et sont issues des études communautaires de référence conduites dans les filières poule pondeuse, poulet de chair, dinde et porc et du plan national de surveillance de l'antibiorésistance pour le bœuf.

Informations disponibles¹

A partir de ces deux sources, diverses informations ont été utilisées : le nombre d'unités épidémiologiques testées (bandes ou carcasses), le nombre d'unités positives ainsi que le ou les sérotypes enregistrés pour chacune des unités épidémiologiques (1 souche retenue par sérotype et par unité²) et, pour les volailles, la capacité du bâtiment abritant la bande échantillonnée. Pour les porcs, les données utilisées sont issues des ganglions mésentériques qui reflètent la contamination en élevage. Un résumé des protocoles de ces études est disponible en annexes 2 et 3.

¹ Seules les informations utilisées pour l'adaptation des données sont indiquées

² Plusieurs mises en culture sont réalisées à partir des échantillons ce qui peut mener à l'isolement de plusieurs sérotypes

Ces données concernent, pour les données actives (table 2.1), 519 élevages de poules pondeuses dont 93 positifs, 371 élevages de poulets de chair dont 32 positifs, 331 élevages de dindes dont 51 positifs, 1 166 carcasses de porcs dont 212 positives et 334 carcasses de bovins dont 8 positives.

	Unités épidémiologiques testées	Unités épidémiologiques positives	Souches collectées*	Année
Pondeuses	519	93	115	2004-2005
Poulets de chair	371	32	34	2005-2006
Dindes	331	51	52	2006-2007
Porcs	1 166	212	212	2006-2007
Bovins	334	8	8	2005

Table 2.1 : Données de surveillance active

* Si 2 sérotypes sont associés à un même prélèvement, 2 souches sont comptabilisées pour ce prélèvement

Ces données étant basées sur un échantillonnage de troupeaux ou de carcasses représentatifs de la production française, elles donnent ainsi accès à des prévalences par sérotype nationales et représentatives. Cependant, le nombre de filières concernées est limité en raison du coût de ce type de surveillance et selon les filières, ces données ne sont pas toujours disponibles pour l'année de référence 2005. Ainsi, pour les poulets de chair, dindes et porcs, les résultats utilisés dans les travaux d'attribution seront ceux de 2006 ou 2007.

2.1.2 Données de surveillance passive

Sources de données

Les données de surveillance passive sont issues de la base de données du Réseau *Salmonella* de l'AFSSA.

Informations disponibles¹

Les données communiquées par le Réseau *Salmonella* concernent des souches accompagnées des informations suivantes : sérotype, origine de la souche (filière et libellé détaillé, ex : poulet jaune), type d'échantillon, point d'échantillonnage (élevage ou abattoir), statut clinique. Pour cette étude, 9 sources ont été retenues : les 5 filières précédemment

¹ Seules les informations utilisées pour l'adaptation des données sont indiquées

citées ainsi que le canard, les produits de la mer, le mouton et les volailles autres (pintade, caille, oie, faisan, pigeon).

En 2005, le Réseau *Salmonella* a centralisé 9 914 souches pour le point de mesure élevage et 1 503 souches pour le point de mesure abattoir et produits animaux pour les filières que nous considérons. Elles sont réparties en 4 à 80 sérotypes selon les filières et points de mesure. Le plus grand nombre de souches est centralisé pour le canard en élevage (n=3 328), le plus faible, pour les produits ovins (n=9). Ces données brutes sont détaillées dans la table 2.2.

	Elevage		Abattoir	
	Nb souches	Nb sérotypes	Nb souches	Nb sérotypes
Poules pondeuses (œufs et ovoproduits)	2812	80	61	10
Poulets de chair			162	27
Dindes	851	38	39	9
Porcs	118	18	362	29
Bovins	447	34	292	45
Ovins	135	16	9	4
Produits de la mer	-	-	44	22
Canards	3328	41	151	16
Autres volailles	2223	Non connu	383	Non connu

Table 2.2 : Données de surveillance passive

Ces données étant basées sur l'envoi volontaire de souches par les laboratoires d'analyse, elles concernent simplement des répartitions par sérotypes (et non des prévalences), nationales, mais la représentativité est dans ce cadre inconnue. En revanche, elles sont disponibles chaque année pour l'ensemble des filières et permettent d'avoir accès à un grand nombre de filières animales, à la fois au point de mesure élevage et au point de mesure abattoir.

2.2 Adaptation des données pour l'attribution

2.2.1 Données de surveillance active

Deux indicateurs sont calculés : la prévalence, globale et par sérotype, et la répartition par sérotype.

Pour les volailles, afin de représenter la production de carcasses ou d'œufs de consommation et en raison de la variabilité des tailles de troupeaux (400 à 20 200 animaux pour les poules pondeuses), les prévalences globales et par sérotype sont pondérées par la taille de bande. La répartition des sérotypes est calculée sans pondération.

2.2.2 Données de surveillance passive

A partir des données concernant les souches d'abattoir et les souches d'élevage (animaux et environnement d'élevage) collectées en 2005, les souches pour lesquelles le libellé était imprécis ont été systématiquement écartées. Au point de mesure élevage, les souches issues d'élevages de reproduction et de cas cliniques ont été exclues. Au point de mesure abattoir, seules les souches issues de carcasses et les morceaux complets (peu découpés, avec la peau) ont été retenues. Pour les filières ponte et produits de la mer, le point de mesure « abattoir » intègre des données concernant, respectivement, les œufs coquilles et les poissons, crustacés ou coquillages non préparés.

Les données d'élevage et d'abattoir ont été regroupées pour chacune des filières à l'exception des bovins pour lesquels une surreprésentation de sérotypes pathogènes a été détectée en élevage, seules les données d'abattoir ont donc été retenues pour cette filière. Combiner les données d'élevage et d'abattoir permet de disposer d'un jeu de données plus important par espèce et d'intégrer les filières pour lesquelles aucune donnée amont n'est disponible à l'exemple des produits de la mer.

2.3 Résultats

2.3.1 Données de surveillance active

Les prévalences et répartitions par filière sont présentées dans l'annexe 4. Les prévalences globales se répartissent comme suit : poule pondeuse (30,8%), porc (18,8%), dinde (13,8%), le poulet de chair (8,2%) et bovin (2,4%). Une forte prévalence est donc observée chez la poule pondeuse, alors que pour le bovin elle est très faible. Un total de 31 sérotypes est observé chez les pondeuses, dont les plus fréquents en prévalence sont Typhimurium (6,9%), Infantis (5,5%) et Anatum (5,3%), mais aucun sérotype ne semble prédominant. Chez le porc, parmi les 28 sérotypes observés, Typhimurium et Derby (7,6% et 6,5%) ont une prévalence très supérieure aux autres sérotypes dont la prévalence est inférieure à 1%. En filière dinde, 3 sérotypes, parmi les 15 observés, ont une prévalence supérieure à 1% : Derby (3,6%), Typhimurium (2,6%) et Mbandaka (1,6%). Chez le poulet de chair, seuls 3 sérotypes sur 23 sont au-dessus de 1% : Hadar (1,6%), Virchow (1,5%) et Montevideo (1,3%). Enfin, le sérotype majoritaire chez le bovin est Mbandaka (1,5%), mais seuls 3 sérotypes sont observés et concernent uniquement 7 souches. On notera que sur les 5 sérotypes cibles des plans de contrôle européens (Typhimurium, Enteritidis, Infantis, Hadar, Virchow), seul Enteritidis n'apparaît jamais parmi les trois sérotypes les plus fréquents pour une filière.

Les distributions de sérotypes, issues des prévalences sont présentées sur le graphe ci-dessous (figure 2.3). Ces distributions sont significativement différentes d'une espèce à l'autre (test exact de Fisher).

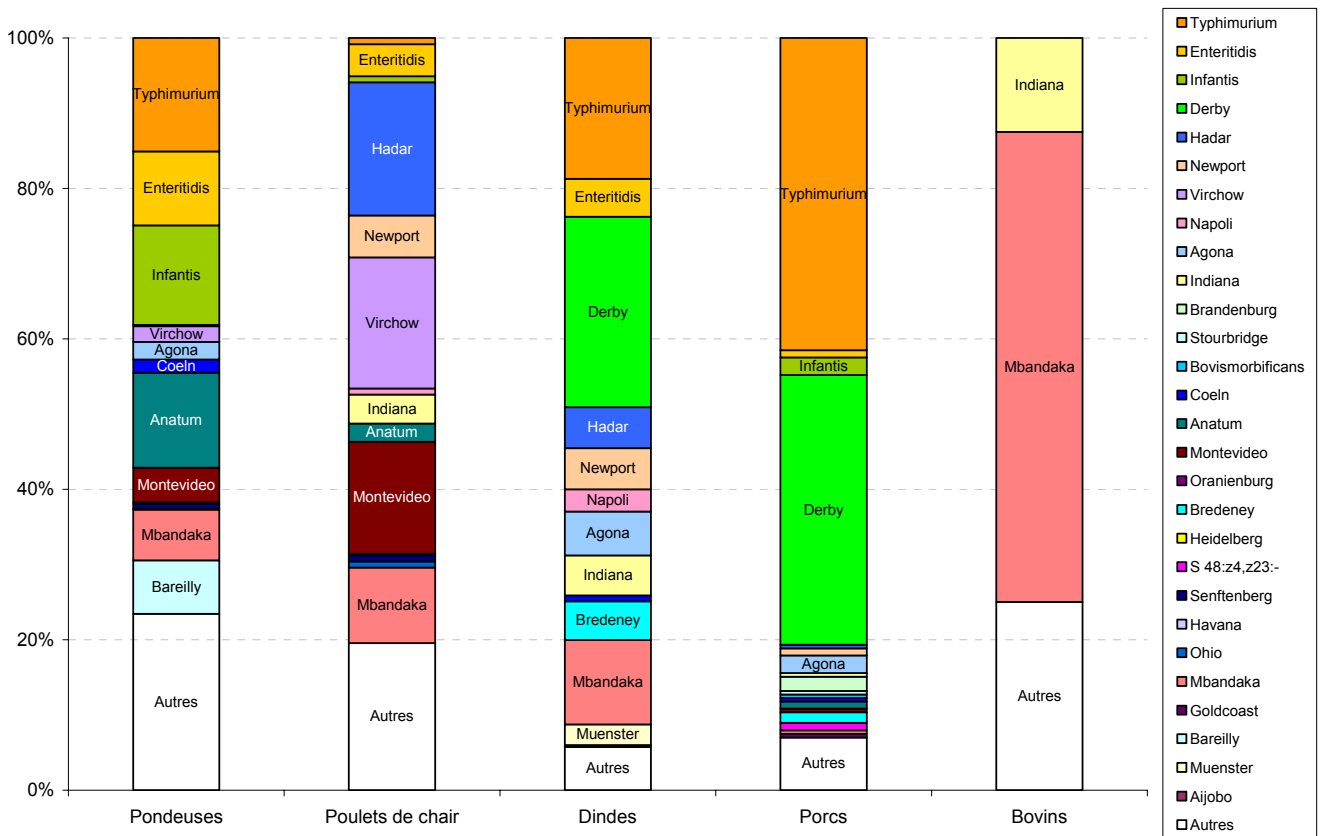


Figure 2.3 : Répartition des souches par sérotype en fonction de la filière pour la surveillance active

2.3.2 Données de surveillance passive

Les répartitions par filière des données passives poolées pour l'élevage et l'abattoir, sont présentées dans l'annexe 5 et la figure 2.4. Les volailles sont caractérisées par la présence d'une grande variété de sérotypes. Pour ces espèces aviaires, quatre sérotypes semblent prédominants quelle que soit la filière : Enteritidis, Typhimurium, Indiana et dans une moindre mesure, Senftenberg. Les sérotypes Typhimurium et Derby sont majoritaires chez le porc et les bovins, alors que le sérotype S 61:k:1,5,7 marque la filière ovine pour laquelle il représente près de 60% des souches. Enfin, les sérotypes les plus fréquents pour les produits de la mer sont Weltevreden et Typhimurium. On note ici l'ubiquité du sérotype Typhimurium présent de façon importante dans l'ensemble des filières.

Enfin, malgré les traits communs pour certaines espèces, et notamment les espèces aviaires, les répartitions sont significativement différentes d'une espèce à l'autre (test exact de Fisher).

On obtient ainsi deux jeux de données concernant les sources animales dont les caractéristiques sont différentes en termes d'indicateur disponible, de représentativité, de

durée de surveillance, de nombre de sources, et de coïncidence temporelle avec les données humaines.

Pour chacune des filières communes considérées (poules pondeuses, poulets de chair, dindes, porcs, bovins), selon le mode de surveillance, les résultats en terme de répartition des types sont significativement différents (test exact de Fisher), même si des corrélations de rangs significatives apparaissent pour les filières ponte et poulet de chair (test des rangs de Spearman). On obtient ainsi pour chacune de ces filières deux images différentes dont l'impact sur les résultats du modèle d'attribution sera étudié dans la deuxième partie de cette thèse.

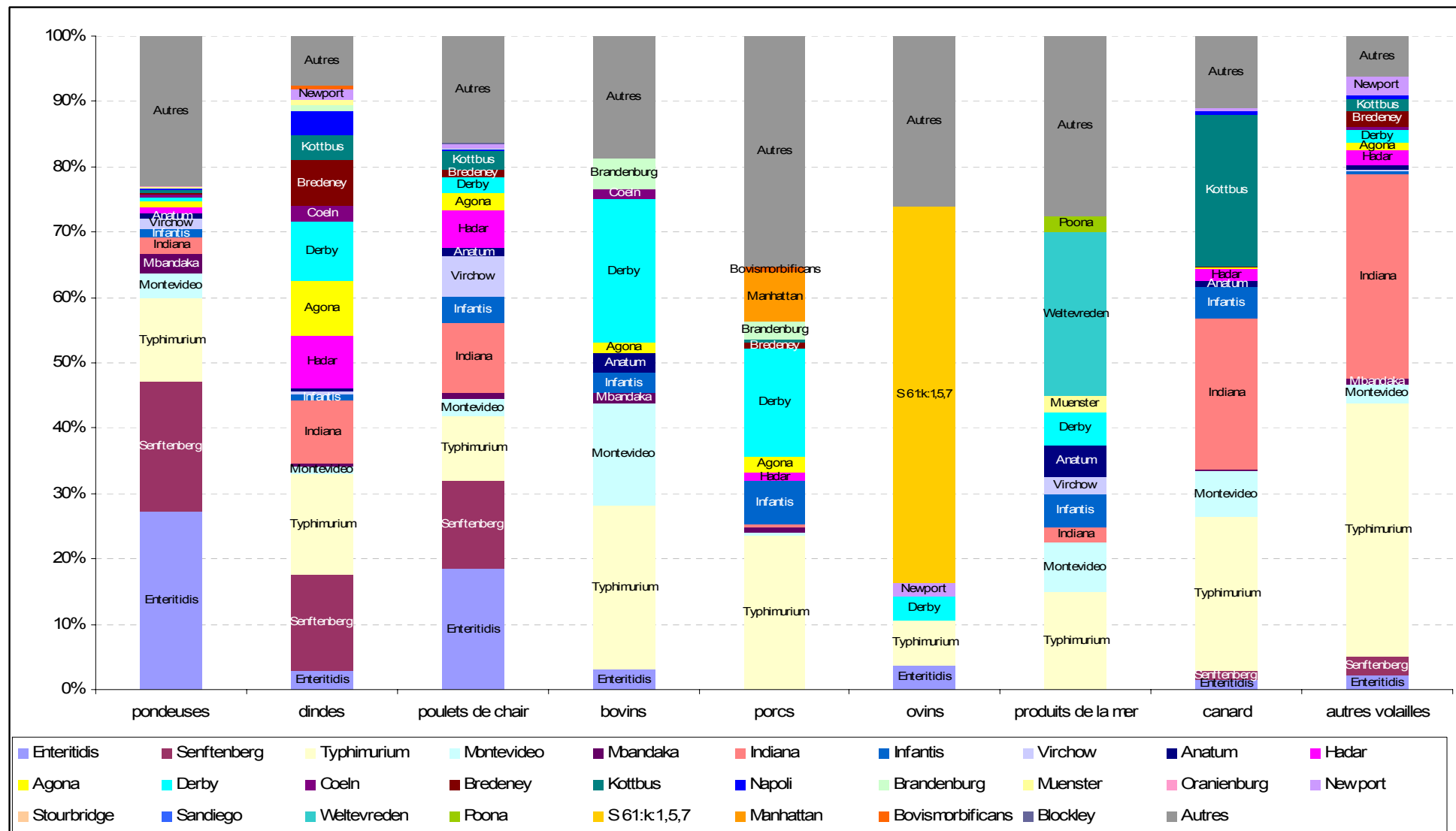


Figure 2.4 : Répartition des souches par sérotype en fonction de la filière pour la surveillance passive

3 Sous typage

Comme vu précédemment, les sérotypes Typhimurium et Enteritidis sont fortement majoritaires parmi les cas humains (Cf. § 1). De plus, le sérotype Typhimurium est ubiquitaire dans l'ensemble des filières animales et il en va de même pour le sérotype Enteritidis dans les filières aviaires. Le sérotypage n'est donc pas suffisamment discriminant. Il est ainsi nécessaire de sous-typer les souches de ces deux sérotypes afin d'affiner l'attribution. Pour ce faire, il faut disposer d'un outil de typage commun aux laboratoires analysant les souches humaines et animales et utilisé en routine afin que des résultats de sous-typage comparables soient disponibles pour toutes les souches incluses dans les travaux. L'antibiotypie a donc été utilisée en premier lieu. Il s'agit d'une méthode de routine et réalisée de façon harmonisée par les laboratoires impliqués dans la surveillance tant humaine qu'animale. Pour conforter les résultats obtenus sur la base de cette méthode phénotypique, nous avons souhaité disposer d'une méthode génotypique. Le typage par Electrophorèse en Champ Pulsé (PFGE), ou pulsotypage, d'un échantillon de souches humaines et de l'ensemble des souches animales issues de la surveillance active a donc été réalisé dans le cadre de ce projet, par le CNR pour les souches humaines et par l'équipe du Réseau Salmonella en tant que LNR Salmonelles associé, pour les souches animales. Cette méthode est habituellement utilisée en cas d'investigation d'épidémie.

3.1 Sous-typage par PFGE

3.1.1 Méthode

Le pulsotypage a été réalisé selon le protocole pulsenet (Peters, Maguire et al. 2003) avec une seule enzyme de restriction (Xba1). Les profils génomiques obtenus pour les souches humaines et les souches animales ont été rassemblés pour analyse visuelle commune afin d'établir des correspondances. Les résultats portent pour Typhimurium sur 56 souches humaines et 97 souches animales et pour Enteritidis sur 48 souches humaines et 9 souches animales. En raison du coût de ces analyses et de leur caractère exploratoire, elles ont porté pour les souches humaines sur un échantillon de taille réduite (50 souches pour chaque sérotype sur plus de 3000 cas). Pour les souches animales, il a été choisi de sous-typer

l'ensemble des souches issues de la surveillance active (106 souches au total) et de ne pas pulsotyper les souches issues de la surveillance passive.

3.1.2 Résultats

Pour les souches de sérotype Typhimurium, sur les 24 profils observés pour les souches humaines et les 29 profils observés pour les souches animales, 5 profils communs ont été identifiés permettant potentiellement d'attribuer 55% des souches humaines ainsi typées. Concernant Enteritidis, un seul profil commun est détecté parmi les 20 profils humains et 5 profils animaux. Il représente 37,5% des souches humaines et 22,2% des souches animales. La PFGE apparaît donc insuffisamment discriminante pour être utilisée dans le modèle d'attribution, notamment pour les souches de sérotype Enteritidis.

3.2 Sous-typage par antibiotypie

3.2.1 Méthode

Profils d'antibiorésistance

La sensibilité aux antibiotiques est déterminée par la méthode des disques de diffusion sur milieu de Mueller Hinton, selon les recommandations du Comité de l'Antibiogramme de la Société Française de Microbiologie (CA-SFM). Les diamètres critiques utilisés sont ceux recommandés par le CA-SFM. Cette méthode est considérée comme fiable, répétable, économique et parfaitement adaptée pour le sous-typage de bactéries (Berge, Atwill et al. 2003). La définition des souches résistantes, intermédiaires ou sensibles est une définition clinique. Malgré nos objectifs épidémiologiques, nous avons choisi de conserver cette méthode de catégorisation des résistances qui est la méthode de routine pour les laboratoires impliqués (CNR et Réseau Salmonella).

Les profils sont définis sur la base des antibiotiques communément testés par le CNR et le Réseau Salmonella : amoxicilline-ampicilline (Amx), chloramphénicol (C), ceftazidime (Caz), gentamicine (Gm), kanamycine (K), acide Nalidixique (Na), streptomycine (S), sulfamides (SSS), sulfamides-triméthoprimine (Sxt) et tétracycline (Te). Les souches sont classées en sensibles et non sensibles (intermédiaires et résistantes). Cette dernière catégorie sera désignée par abus de langage comme résistante par la suite. La seule exception à cette règle

est le chloramphénicol pour lequel, pour des raisons de cohérence génétique, les souches intermédiaires sont considérées comme sensibles.

Classification des souches

Un grand nombre de profils est observé pour chacun des sérotypes, 17 pour Enteritidis et 42 pour Typhimurium. Il est donc nécessaire de les regrouper en clusters dont les effectifs seront optimisés. La classification porte sur les souches résistantes à au moins un antibiotique. Une analyse en correspondances multiples (ACM), adaptée aux données qualitatives, a été appliquée aux résultats des antibiogrammes et combinée à une classification de type mixte qui permet de créer des clusters de souches à partir des coordonnées factorielles issues de l'ACM. Les classifications optimales ont été choisies conformément aux recommandations de Berge *et al*, selon les effectifs et l'homogénéité des clusters proposés ainsi que la variabilité prise en compte par la partition (Berge, Atwill et al. 2003). Le détail de la technique est présenté en annexe 6.

3.2.2 Résultats pour Enteritidis

Les souches d'Enteritidis (n=202) sont majoritairement sensibles à tous les antibiotiques testés (n=137 soit 68%). La multirésistance (résistance à au moins 4 antibiotiques) n'est observée que pour une souche humaine (figure 2.5). Les souches résistantes (n=65) sont principalement caractérisées par leur résistance à l'acide nalidixique (n=39, 60% des souches résistantes à au moins un antibiotique), aucune n'est résistante au ceftazidime. Au total, 16 profils différents sont observés, le profil majoritaire étant la résistance simple à l'acide nalidixique (n=31, 15% de l'ensemble des souches).

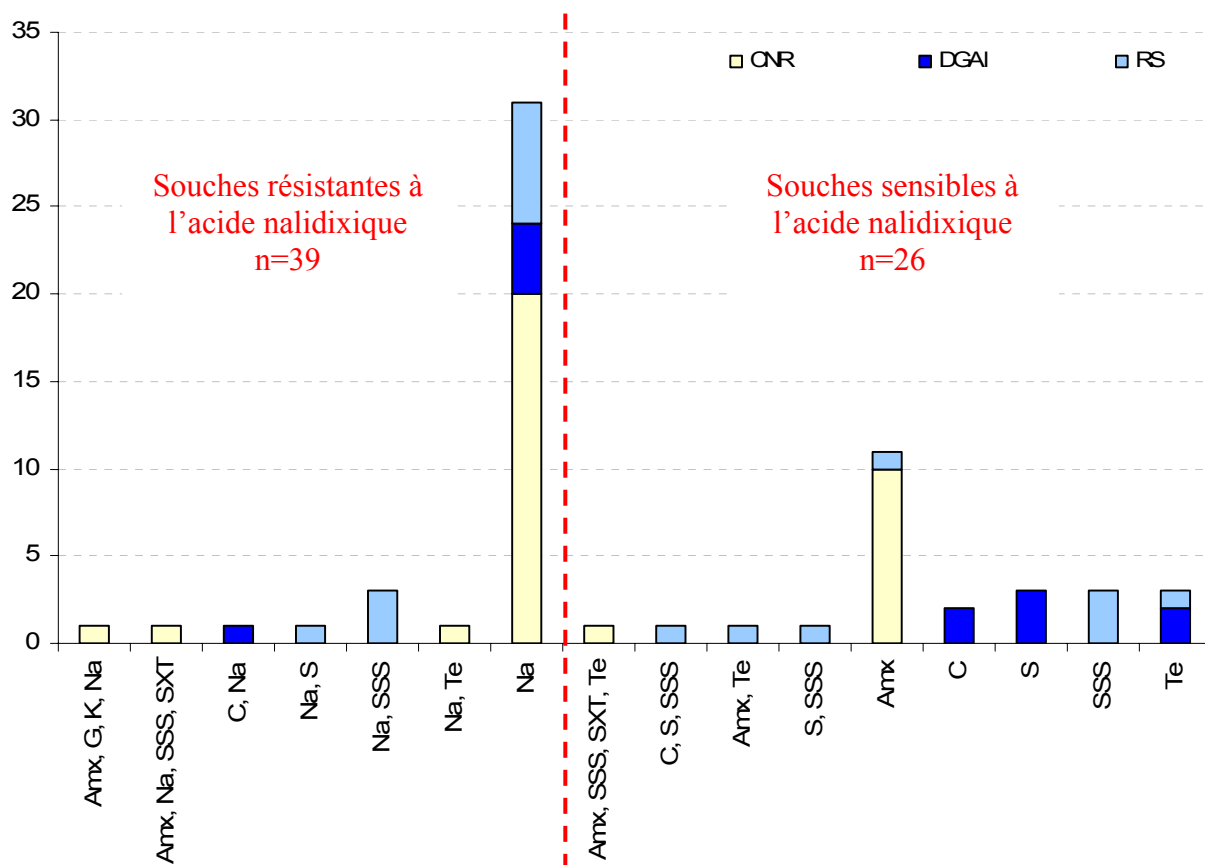


Figure 2.5 : Répartition des souches de sérotype Enteritidis par profil et par source de données

La résistance à l'acide nalidixique est chromosomique (Bager and Helmuth 2001; Cloeckert and Chaslus-Dancla 2001), donc stable (en comparaison avec les résistances portées et véhiculées par les plasmides, éléments génétiques mobiles pouvant être transmis d'une bactérie à l'autre). Sur les conseils du CNR, les souches d'Enteritidis résistantes à au moins un antibiotique ont été divisées en un groupe résistant à l'acide nalidixique et un groupe sensible à cet antibiotique. La classification a été réalisée séparément sur ces deux groupes. Pour le groupe résistant à l'acide nalidixique (NaR, n=39), un groupe NaR simple (SE1) et un groupe associant la résistance à l'acide nalidixique à au moins une autre résistance (SE2) ont été identifiés. Les deux classes rassemblent des souches humaines et animales.

La partition retenue pour le groupe sensible à l'acide nalidixique est composée de 6 clusters. La classe 1 (SE3) est caractérisée par la résistance à l'amoxicilline, la classe 2 (SE4), par la résistance à la tétracycline, la classe 4 (SE6), par la résistance aux sulfamides, la classe 5 (SE7), par la résistance à la streptomycine et la classe 6 (SE8) par la résistance au chloramphénicol. La classe 3 (SE5) est constituée par la souche humaine multi-résistante (AmxSssSxtTe). Seule la classe 1 rassemble à la fois des souches humaines et des souches animales.

Parmi les 9 sous-types ainsi définis, seuls les cas humains relatifs aux trois premiers sous-types et aux Enteritidis multi-sensibles (SE-multiS) seront attribuables à une source animale

Groupe	Caractéristique principale	Origine	Nb de cas humains associés*
SE multiS	Multi-sensibles	Homme et animal	2092
SE1	Résistance à l'acide nalidixique seule	Homme et animal	615
SE2	Résistance à l'acide nalidixique associée à une autre résistance	Homme et animal	92
SE3	Résistance à l'amoxicilline	Homme et animal	308
SE4	Résistance à la tétracycline	Animal	-
SE5	Multirésistance	Homme	31
SE6	Résistance aux sulfamides	Animal	-
SE7	Résistance à la streptomycine	Animal	-
SE8	Résistance au chloramphénicol	Animal	-

Table 2.6 : Caractéristiques des sous-types pour Enteritidis

* au sein d'un sérotype, le nombre de cas pour chaque sous-type est obtenu en appliquant la proportion observée parmi les souches typées à l'ensemble des souches d'Enteritidis ou de Typhimurium

3.2.3 Résultats pour Typhimurium

Les souches de Typhimurium (n=342) sont, pour la moitié, multirésistantes (n=168, 49%) (figure 2.6). Le phénotype multi-sensible (sensible à l'ensemble des antibiotiques testés) n'est observé que chez les souches humaines et chez une seule souche animale (souche de poulet de chair issue de la surveillance passive). L'absence d'autre souche de Typhimurium multi-sensible parmi les souches animales peut être associée à un problème d'interprétation d'antibiogramme, lié plus particulièrement à la streptomycine. En effet, cet antibiotique, dont la résistance simple est fréquente parmi les souches animales, pose des problèmes d'interprétation (Doran, NiChulain et al. 2006). Les valeurs des diamètres et breakpoints critiques sont d'ailleurs en cours de révision au niveau européen dans le cadre de Eucast (European Committee on Antimicrobial Susceptibility Testing). De plus, la lecture dans le cas des souches humaines est une lecture interprétée. L'harmonisation de l'interprétation de la streptomycine nécessiterait donc d'avoir l'ensemble des données d'antibiorésistance sous forme de diamètres d'inhibition. En l'absence de ces informations pour l'ensemble des souches, la souche issue du poulet de chair de la base de données passive a été exclue. Les 1 115 cas humains correspondants ne seront donc pas attribuables dans le cadre d'une démarche d'attribution.

Cinq résistances sont fréquemment observées, les résistances à l'amoxicilline, au chloramphénicol, à la streptomycine, aux sulfamides et à la tétracycline (Figure 2.8). Ces

antibiotiques correspondent à la pentarésistance chromosomique ACSSSTe associée chez *Typhimurium* le plus souvent au lysotype DT104 (des souches de DT 120 porteuses de la même résistance, mais plasmidique, sont cependant apparues ces dernières années). Le profil le plus fréquent (38,6% des souches) est en effet la penta-résistance [Amx, C, S, SSS, Te]. Un autre profil est fréquemment observé parmi les souches animales et concerne 18,4% de l'ensemble des souches, il s'agit de la résistance simple à la streptomycine. Comme pour *Enteritidis*, aucune souche n'est résistante au ceftazidime.

La partition choisie conformément aux critères de Berge est en 8 classes. La première classe rassemble les souches penta-résistantes. Les classes 2, 3 et 4 rassemblent des souches dont le profil est proche de la penta-résistance avec quelques résistances en moins ou en plus : La classe 2 avec une association systématique à la résistance à l'acide nalidixique, la classe 3 avec une association systématique à la résistance aux sulfamides-triméthoprimine, la classe 4 (uniquement des souches animales) avec une association systématique à la kanamycine. La classe 5 rassemble l'ensemble des souches sensibles à la streptomycine, il s'agit uniquement de souches humaines. La classe 6 correspond à des souches résistantes à la streptomycine et à la tétracycline avec quelques résistances à l'amoxicilline et aux sulfamides. La classe 7 est caractérisée par une résistance à la streptomycine associée pour quelques souches à la résistance aux sulfamides. Enfin, la classe 8 est caractérisée par une double résistance systématique à l'acide nalidixique et à la streptomycine. Il est à noter que les classes 2 à 6 et 8 ne sont pas homogènes. Les classes 4 et 8 contiennent exclusivement des souches animales et la classe 5, exclusivement des souches humaines. Ces souches ne pourront donc pas être attribuées dans le cadre d'une attribution par typage microbiologique. In fine, 9 clusters sont déterminés : un cluster multi-sensible (STmultiS) et 8 clusters de souches résistantes à au moins un antibiotique (ST1 à ST8).

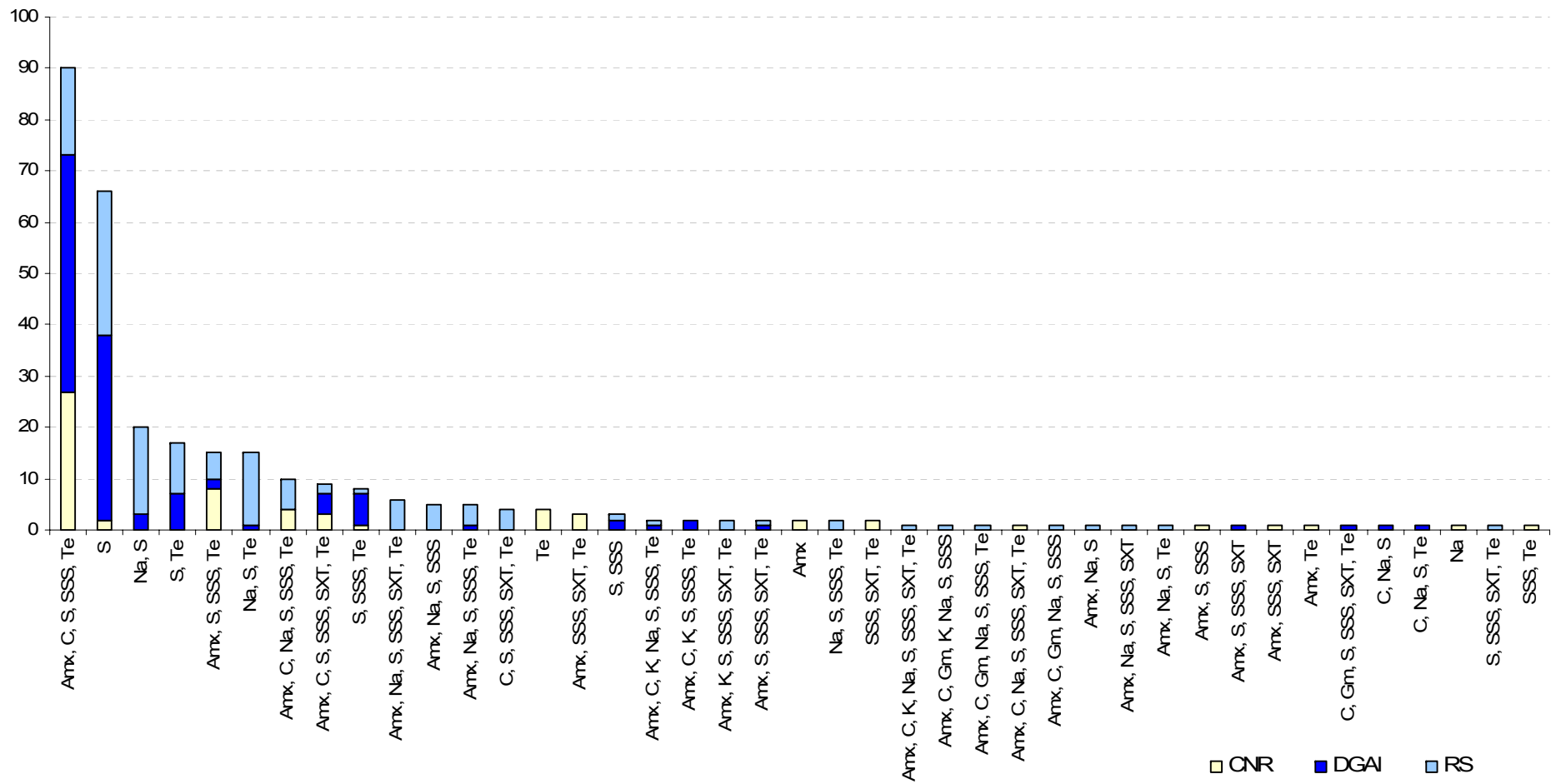


Figure 2.6 : Répartition des souches de sérotype Typhimurium par profil et par source de données

Groupe	Caractéristique principale	Origine	Nb de cas humains associés
ST multiS	Multi-sensibles	Homme et animal	1115
ST1	Pentarésistance AmxCSSsTe	Homme et animal	1038
ST2	AmxCSSsTe et résistance à l'acide nalidixique	Homme et animal	154
ST3	AmxCSSsTe et résistance aux sulfamides-triméthopime	Homme et animal	154
ST4	AmxCSSsTe et résistance à la kanamycine	Animal	-
ST5	Sensibilité à la streptomycine	Homme	615
ST6	Résistance à la streptomycine et à la tétracycline	Homme et animal	384
ST7	Résistance à la streptomycine	Homme et animal	77
ST8	Résistance à la streptomycine et à l'acide nalidixique	Animal	-

Table 2.7 : Caractéristiques des sous-types déterminés par ACM pour Typhimurium

4 Données de consommation

On cherche ici à obtenir les données de consommation des produits issus des filières considérées pour l'année de référence 2005 en population française.

4.1 Données disponibles

Deux sources de données ont été utilisées, leurs caractéristiques sont indiquées dans la table 2.8.

Source	Type de données	Date de l'étude	Description brève	Représentativité
Office de l'élevage OFIMER	Bilan de disponibilité par filière en tec	Données annuelles	Estimation des quantités consommées au niveau national par bilan entre production, importation et exportation, agrégées au niveau filière	Exhaustivité sur le territoire national
INCA (AFSSA)	Consommation réelle	Etude ponctuelle en 1999	Relevé des consommations détaillé sur 1 semaine + questionnaire général précis 4 vagues d'enquête (prise en compte de la saisonnalité) Taille des portions évaluée par rapport à des photos	3 ans et + Recrutement national aléatoire selon la méthode des quotas n = 3003

Table 2.8 : données de consommation

tec : tonnes équivalent carcasse ou tonnes équivalent coquilles

Ainsi les données sont disponibles à deux points de la chaîne agro-alimentaire : disponibilité sur le marché et consommation réelle.

4.2 Construction de l'indicateur

Les données issues de l'enquête INCA (enquêtes individuelles nationales de consommations alimentaires) permettent d'avoir accès à la consommation réelle au niveau national. Cependant, ces données ont été collectées en 1999¹. Comme la consommation de produits animaux fluctue de façon importante d'une année à l'autre (figure 2.7) selon des tendances de fond (substitution des viandes rouges par les viandes blanches) et au gré des crises sanitaires (ESB, dioxine), il était nécessaire d'actualiser les valeurs obtenues pour pouvoir les utiliser avec des données de prévalences et de cas humains de 2005 (Combris 1997). L'actualisation a été réalisée grâce à un indice en base 100 en 1999 (l'année d'INCA), constitué sur la base des données de disponibilité sur le marché, en raison de leur facilité d'accès et de leur publication annuelle.

¹ Une nouvelle enquête a été réalisée en 2006, mais ses résultats n'étaient pas encore disponibles dans le cadre du projet.

4.2.1 Indice en base 100

Les données issues de l'office de l'élevage sont disponibles en ligne (http://www.office-elevage.fr/publications/publication_marche.htm) et permettent d'avoir accès aux filières suivantes : volailles, œufs, porc, bovins, ovins/caprins. L'indice volaille a été utilisé de façon commune pour l'ensemble des volailles de chair (poulet de chair, dinde, canard, autres volailles). Concernant les produits de la mer, les données utilisées sont issues de l'OFIMER (http://www.ofimer.fr/Pages/filiere/Publications_ObsEco.html#chiffresClesFr). L'indice en base 100 obtenu est présenté ci-dessous (Cf figure 2.7).

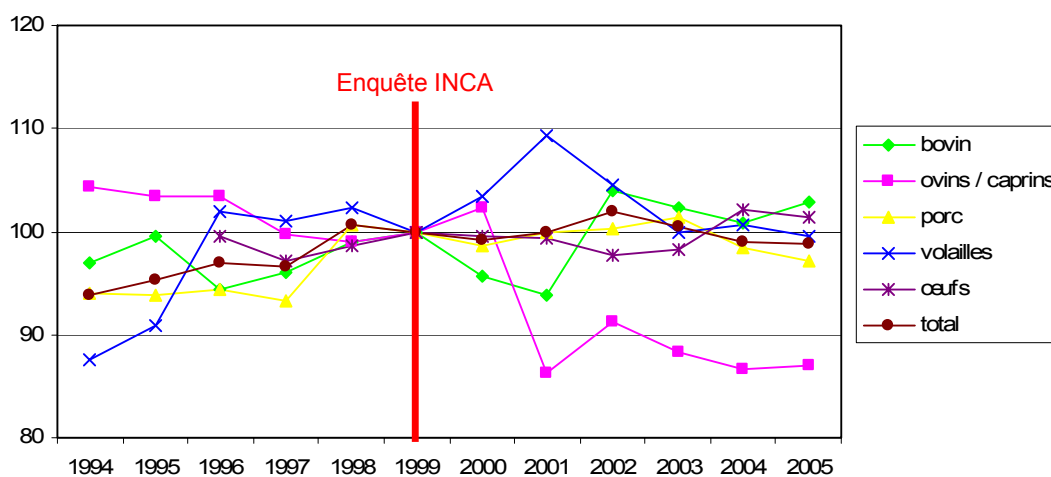


Figure 2.7 : Disponibilité sur le marché des produits animaux sous forme d'un indice en base 100, année de référence 1999, données OFIVAL

Par rapport à 1999, les consommations d'œufs et de bovin bénéficient d'une légère hausse (+1,5% et +3%), alors que les autres consommations subissent des baisses : volailles (-0,5%), le porc (-3%) et ovins (-13%). En ce qui concerne les produits de la mer, non représentés ici, l'évolution entre 1999 et 2005 se traduit par une forte hausse +21%.

4.2.2 Obtention des données de consommation à partir d'INCA

Afin d'obtenir des données agrégées par filière (bovin, œufs, etc...), il a été nécessaire de sélectionner pour chaque filière les aliments simples ou composés s'y rapportant parmi les aliments reportés par les enquêtés. En effet, dans cette base, les consommations sont relevées de façon très détaillée, et pour une même filière, de nombreux types d'aliments, souvent composites sont enregistrés (Volatier 2000). Pour les œufs, seuls les œufs coquille

ou les préparations à base d'œufs frais (omelette, mousse au chocolat, ...) ont été retenus. Pour les viandes, le cinquième quartier (abats, tripes, ...) a été exclu. Une fois la sélection réalisée, la part de chaque type de viande dans la composition des aliments composés a été évaluée. Pour ce faire, une sélection de recettes est réalisée à partir de la base de recettes des aliments de l'enquête INCA. Pour chacun des produits composites, les proportions respectives de bœuf, porc, etc. ont été calculées sur la base de ces recettes. Enfin, les quantités consommées par personne et par repas ont été sommées pour obtenir une consommation globale. Les résultats obtenus sont les suivants :

Catégories d'aliments	Ensemble de la population			Seuls consommateurs			
	Moyenne	Ecart type	95è perc	Taux de consommateurs	Moyenne	Ecart type	95è perc
	g/j	g/j	g/j	%	g/j	g/j	g/j
Œufs et aliments à base d'œufs	27,84	27,86	80,71	80,05	34,77	26,99	85,71
Bovin et aliments à base de bovin	47,3	37,13	116,86	91,72	51,56	35,82	118,43
Ovin et aliments à base d'ovine	8,31	16,07	39,21	34,46	24,11	19,19	59,81
Poulet et aliments à base de poulet	29,84	38,48	106,43	63,91	46,7	39,12	115
Dinde et aliments à base de dinde	6,26	13,22	29,14	30,05	20,84	16,66	50,71
Porc et aliments à base de porc	57,69	41,81	140	97,76	59,01	41,35	140,15
Produits de la mer	32,82	33,7	95,71	83,04	39,52	33,21	99,29
Canard	1,86	7,44	15,00	8,34	22,24	14,51	57,14
Autres volailles*	1,71	8,69	13,57	5,50	31,21	21,39	64,29

Table 2.9 : Données de consommation chez les adultes de 15 ans et plus (Source : enquête INCA 1999)

*Caille, faisan, pigeon

Guide de lecture : les adultes consomment en moyenne 27,8 g par jour d'œufs ou d'aliments à base d'œufs. Les 5% plus forts consommateurs en consomment plus de 80,7 g/jour. Sur une semaine de consommation, 80% de la population de 15 ans et plus a consommé des œufs ou des produits dérivés: pour ces seuls consommateurs, la consommation moyenne s'établit à 34,8 g/jour. Les taux de consommateurs sont donnés pour la semaine de l'enquête

Catégories d'aliments	Ensemble de la population			Seuls consommateurs			
	Moyenne	Ecart type	95è perc	Taux de consommateurs	Moyenne	Ecart type	95è perc
	g/j	g/j	g/j	%	g/j	g/j	g/j
Œufs et aliments à base d'œufs	21,66	23,78	65,71	75,74	28,61	23,41	74,29
Bovin et aliments à base de bovin	41,22	30,6	95,71	96,27	42,81	30,08	97
Ovin et aliments à base d'ovin	4,82	11,9	23,74	29,57	16,3	17,09	39,91
Poulet et aliments à base de poulet	20,27	28,66	71,86	65,23	31,07	30,39	87,43
Dinde et aliments à base de dinde	6,29	10,76	27,86	39,69	15,85	11,85	40,43
Porc et aliments à base de porc	42,71	28,3	95,04	98,13	43,53	27,94	95,57
Produits de la mer	19,73	20,41	57,14	80,55	24,5	20,02	62,64
Canard	1,30	6,68	7,86	6,19	20,99	17,67	50,00
Autres volailles*	0,67	4,73	0,00	2,85	23,57	15,9	44,29

Table 2.10 : Données de consommation chez les enfants de 3 à 14 ans (Source : enquête INCA 1999)

*Caille, faisán, pigeon

Source : enquête INCA 1999

Les données de consommation moyenne sont ensuite extrapolées à la population entière sur la base des données INSEE du 01/01/1999 (séparément pour les plus de 15 ans et les moins de 15 ans). Enfin, les quantités obtenues sont actualisées grâce à l'indice en base 100 présenté précédemment. In fine, les données de consommation utilisées dans le modèle sont les suivantes.

Filière	INCA 1999 (t)	Indice en base 100	INCA 2005 (t)
Œuf	81123,7	101,5	82301,0
Poulet	85254,2	99,5	84841,7
Dinde	19059,1	99,5	18966,9
Porc	166851,9	97,1	161971,0
Bœuf	140375,2	102,9	144503,9
Mouton	23266,8	87,0	20232,0
Produits de la mer	92290,7	121,1	111764,0
Canard	5335,1	99,5	5309,3
Autres volailles	4602,3	99,5	4580,0

Table 2.11 : Données de consommation utilisées dans le modèle

5 Intégration des données pour la démarche d'attribution

La spécificité du système de surveillance français est de disposer, pour ce qui concerne les réservoirs animaux, de deux systèmes de surveillance complémentaires et parallèles. Ceci représente une opportunité d'étudier l'impact des caractéristiques de ces deux types de données (actives et passives) sur leur utilisation dans le cadre d'une démarche d'attribution. Deux jeux de données différents sont donc constitués, l'un rassemblant les données humaines et les données animales de surveillance active, et l'autre les mêmes données humaines associées aux données animales de surveillance passive.

5.1 Sélection des sérotypes à inclure pour la modélisation

Globalement, les données obtenues pour l'homme et les filières animales se répartissent en 77 sérotypes non typhiques communs entre l'Homme et au moins l'une des filières animales dans l'un des modes de surveillance.

Pour chacun des jeux de données, les sérotypes considérés pour le modèle d'attribution ont été sélectionnés selon les critères suivants :

- nombre de cas humains liés au sérotype : indicateur de l'importance en termes de santé publique ;
- part maximale représentée par le sérotype (proportion) au sein d'une filière animale : indicateur de l'importance dans les sources ;
- types « spécifiques », c'est-à-dire présents dans une seule filière, nécessaires pour assurer la convergence du modèle d'attribution.

Les règles de décision pour les deux premiers critères sont présentées dans la table 2.12. Tous les sérotypes spécifiques ont été retenus. Ainsi 28 sérotypes pour les données de surveillance active et 37 pour les données de surveillance passive ont été inclus. Les autres sérotypes sont rassemblés dans la catégorie « autres ».

Nb de cas humains	Représentation parmi les filières animales
35 cas ou plus	Tous les sérotypes
30 à 34	Proportion \geq 5% dans une source au moins
5 à 29 cas	Proportion \geq 15% dans une source au moins
Moins de 5 cas	Proportion \geq 20% dans une source au moins

Table 2.12 : Règles de décision pour la sélection des sérotypes

Les sous-types de Typhimurium et Enteritidis sont tous inclus pour la modélisation, les cas associés à des sous-types non communs entre l'Homme et l'animal dans le jeu de données considéré sont reclassés dans la catégorie « Typhimurium-autres » et « Enteritidis-autres ».

5.2 Types spécifiques

Le modèle proposé par Hald *et al*/ nécessite la présence de types spécifiques (i.e. présents dans une seule source) qui servent d'ancrage au modèle. Ainsi, pour le jeu de données actif, 14 types spécifiques sont inclus (table II.13) et 15 pour le jeu de données passif (table II.14). On notera que les filières les plus représentées parmi les types spécifiques, en terme de nombre de types ou de cas humains associés, sont les filières porc et poule pondeuse. En revanche, quel que soit le mode de surveillance, aucun type spécifique au bovin n'est disponible.

Type	Source	Prévalence	Nombre de cas
Heidelberg	Poulet de chair	0,02%	30
Ohio	Poulet de chair	0,07%	12
SE NaR2	Poule pondeuse	0,20%	92
Oranienburg	Poule pondeuse	0,09%	31
Havana	Poule pondeuse	0,04%	16
Bareilly	Poule pondeuse	2,95%	8
ST2	Porc	0,09%	154
Brandenburg	Porc	0,34%	71
Stourbridge	Porc	0,09%	42
Bovismorbificans	Porc	0,09%	40
S 48:z4,z23:-	Porc	0,18%	24
Goldcoast	Porc	0,09%	9
Muenster	Dinde	0,38%	7
Aijobo	Dinde	0,04%	2

Table 2.13 : Types spécifiques associés aux données de surveillance active

Type	Source	Proportion	Nombre de cas
Poona	Produits de la mer	2,5%	21
Weltevreden	Produits de la mer	25%	9
Blockley	Poulet de chair	0,1%	29
SE-NaR2	Poule pondeuse	4,1%	92
SE1	Poule pondeuse	1,3%	308
Oranienburg	Poule pondeuse	0,2%	31
Havana	Poule pondeuse	0,2%	16
Manhattan	Porc	7,5%	81
Corvallis	Porc	0,9%	24
Goldcoast	Porc	0,9%	9
S 61:k:1,5,7	Ovin	57,7%	1
Essen	Dinde	0,1%	2
Sandiego	Canard	0,1%	8
S 4,12:-:-	Autres volailles	0,1%	4
Aberdeen	Autres volailles	0,1%	2

Table 2.14 : Types spécifiques associés aux données de surveillance passive

6 Données sur les denrées importées

La consommation de denrées animales se partage entre production nationale et importations. Dans la perspective d'une utilisation en attribution, il serait donc nécessaire de disposer de données spécifiques concernant les produits importés, afin de distinguer le risque lié aux filières de production nationale, sur lesquelles porteront les mesures de maîtrise du risque, du risque lié aux importations sur lequel il n'y a que peu de prise en dehors des contrôles de contamination.

6.1 Importations et consommation française

Selon les données publiées par l'Office de l'Élevage, les importations représentent 15 à 57% de la consommation intérieure brute selon les filières (Cf table 2.15). L'exposition de la population à ces produits importés n'est donc pas négligeable.

Les prévalences et distributions de sérotypes sont différentes entre les élevages issus de différents pays pour une même filière (2009a), par conséquent la distribution de sérotypes observée en élevage en France est très probablement différente de celle dans les produits importés. C'est d'ailleurs ce qui est observé au Danemark concernant leur propre production et leurs importations (Hald, Vose et al. 2004). Extrapoler les données françaises à l'ensemble

des produits consommés semble donc erroné ; pourtant, intégrer les données concernant la contamination des produits importés est essentiel pour tenir compte de la réelle contamination à laquelle la population est exposée.

Ce modèle d'attribution cherche à faire le lien entre cas humains et prévalences en élevage, notamment afin de mettre en place des mesures dans les élevages et de les évaluer. Il est donc important de discerner la part de cas liée aux produits issus des élevages du territoire, objets d'un éventuel management de santé publique, de celle liée aux produits issus d'importation. Si les importations ne sont pas intégrées comme source dans la démarche d'attribution, les cas qui sont liés à la consommation de produits d'importation seront attribués de façon erronée à l'une des sources considérées, menant à un résultat biaisé. Les données de surveillance concernant les produits importés sont donc explorées.

	Consommation Intérieure Brute (tec)	Part d'importation
Ovins / Caprins	260	56,5%
Bovins	1680	24,3%
Porcs	2157	23,8%
Volailles	1440	19,6%
Œufs	951	14,7%

Table 2.15 : Données Office de l'élevage, 2005

6.2 Sources de données explorées

6.2.1 Surveillance des denrées importées en France

Les denrées importées font, en France, l'objet d'une surveillance dans le cadre de la Food Law. Cette activité est coordonnée par la Mission de Coordination Sanitaire Internationale (MCSI) au sein de la DGAL. Or, cette surveillance est ciblée sur des couples pathogène-matrice estimés à risque, afin d'optimiser les moyens et de cibler les contrôles sur la base d'arguments scientifiques de type analyse quantitative de risque, conformément aux recommandations de la Food Law. Pour *Salmonella*, les contrôles concernent essentiellement les filières avicoles. Puisque nous avons besoin de données couvrant l'ensemble des filières ou tout au moins les 5 principales, ces données sont insuffisantes pour une utilisation en attribution.

6.2.2 Utilisation des données de surveillance européennes

Si nous nous intéressons à la provenance des importations (données en valeur, source DGCCRF) pour 2005 (Cf table 2.16), nous constatons que la majeure partie de ces importations est en provenance de l'Union Européenne.

Filière	Part UE	Principal contributeur autre	
Bovins	95,1%	-	-
Ovins / Caprins	68,3%	Nouvelle Zélande	28,4%
Porcs	96,0%	-	-
Poulets de chair	95,7%	-	-
Dindes	74,1%	Brésil	18,9%
Œufs	96,8%	-	-

Table 2.16 : Origine des importations (Données Douanes 2005, en euros))

Ceci apparaît particulièrement vrai pour les filières bovine, porcine et ovoproduits. Or, dans le cadre de la directive zoonose, la surveillance active des salmonelles est devenue obligatoire pour les poules pondeuses, poulets de chair, dindes et porcs. Nous avons donc cherché à combiner les informations de structure d'importation par pays qui nous ont été fournies par la DGCCRF et les données de prévalence par pays publiées par l'EFSA. Cette démarche n'a pas pu aboutir. En effet, le sous-typage des souches d'Enteritidis et de Typhimurium est indispensable pour pouvoir utiliser les données dans le cadre de notre modèle d'attribution, et ces données ne sont pas disponibles auprès de l'EFSA. Quelques pays ont accepté de communiquer les profils d'antibiorésistance de leurs souches (Pays-Bas, Belgique) mais cela s'est avéré insuffisant. Enfin, pour les filières pour lesquelles des pays hors Union Européenne jouent un rôle majeur dans les importations, comme la Nouvelle-Zélande pour les ovins, utiliser une telle démarche suppose de créer des collaborations sur le plan international qui n'existent pas à l'heure actuelle à notre connaissance.

6.3 Conclusion

Si des informations fiables sont disponibles en France sur la structure des importations par filière et sur la contamination de certaines matrices à risque par *Salmonella*, il n'existe pas de base de données globale représentative de la contamination des denrées importées. Pour pallier ce problème, les importations étant principalement issues de l'Union Européenne, des informations issues des bases de données européennes ont été collectées. Néanmoins, le sous-typage n'ayant pu être obtenu, les données concernant les denrées importées n'ont pas

été intégrées au modèle. En conséquence, les résultats présentés ci-après sont à considérer avec prudence.

Le système de surveillance intégré français des salmonelles permet de rassembler les données indispensables à une démarche d'attribution par typage microbiologique : recensement des cas humains domestiques sporadiques, contamination des sources animales potentielles de ces cas et consommation des aliments issus des sources considérées. L'application harmonisées des outils de sérotypage et de sous-typage permet en outre de disposer de données parfaitement adaptées à une telle démarche.

La constitution de la base de données a reposé sur des collaborations de grande ampleur entre les différents acteurs de la surveillance et a nécessité leur engagement notamment pour ce qui concerne la réalisation du sous-typage. La mise en place de telles collaborations est à l'heure actuelle la clé de voûte de tels projets transversaux, dont les implications se mesurent non seulement en termes de santé publique mais également en termes d'enjeux économiques et de commerce international pour les filières animales. La dynamique développée autour du projet permet d'être optimiste quant à la pérennité de la démarche.

Les données collectées souffrent cependant de certaines insuffisances. L'absence de données sur les importations est un point important. En effet, la part de denrées importées dans la consommation française pouvant représenter jusqu'à 60% de la consommation pour une filière, ne pas tenir compte des importations peut mener à attribuer abusivement des cas aux filières françaises. D'autre part, les données de surveillance active concernant les sources animales présentent pour trois filières majeures (poulets de chair, dindes et porcs) un décalage temporel avec les données humaines. De plus, le sous-typage des souches d'*Enteritidis* et de *Typhimurium* n'a porté chez l'Homme que sur 3% des cas, l'incertitude générée dans les résultats du modèle sera donc potentiellement importante. En revanche, disposer de données de surveillance active et de données de surveillance passive concernant les sources offre la possibilité d'explorer l'utilisation respective des ces données de nature différente dans le modèle. Ainsi, nous avons choisi d'orienter la partie modélisation de la thèse vers l'optimisation de l'outil d'attribution proposé par Hald *et al* en fonction des données disponibles en France plutôt que vers la production de résultats d'attribution en propre.

Deuxième partie

Le modèle d'attribution par typage microbiologique

III

Application de l'attribution dans divers pays européens

IV

Sensibilité du modèle aux priors informatifs

V

Impact de la qualité des données sur les résultats obtenus avec le modèle

Le jeu de données qui a pu être constitué à partir du système de surveillance français, malgré certaines imperfections, couvre les prérequis nécessaires à l'application de la méthode bayésienne d'attribution par typage microbiologique telle que décrite par Hald *et al* (2004). Cette deuxième partie concerne donc l'application de l'outil à un jeu de données considéré comme suffisant pour utiliser le modèle, mais dont les caractéristiques sont différentes du jeu de données danois¹, pour lequel il a été mis au point.

Dans le cadre du réseau d'excellence Med-Vet-Net et plus particulièrement du projet européen d'attribution (« Human illness source attribution ») auquel nous avons participé, différents pays ont appliqué la méthode proposée par les Danois à leurs propres données. Le chapitre III présente les résultats de ces démarches, qui se sont, pour certaines, soldées par des échecs. Le modèle proposé nécessite d'inclure de l'information *a priori* informative pour certains paramètres, ce qui n'est pas conseillé dans une démarche bayésienne. Dans un premier temps, l'impact de cette information *a priori* informative a été étudié. Dans un deuxième temps, une alternative d'information informative reposant sur l'utilisation de l'information contenue dans le jeu de données utilisé a été proposée (chapitre IV). Enfin, le manque de qualité des données ayant été désigné comme possible explication des échecs dans les conclusions de l'article européen, l'impact de la qualité des données sur les résultats d'attribution a été étudié à partir des données françaises (chapitre V). Ce dernier chapitre présente également des résultats d'attribution préliminaires pour la France et permet de dégager les nécessaires améliorations à apporter au jeu de données pour obtenir des résultats d'attribution fiables.

¹ En termes d'indicateur, sources prises en compte, représentativité et outil de typage utilisé

III

Application de l'attribution dans divers pays européens

Troisième article

Attribution de source appliquée à *Salmonella* dans différents
pays européens

Article en préparation pour soumission à Foodborne Pathogens and
Disease

Dans le cadre du projet « Human illness source attribution » du réseau d'excellence Med-Vet-Net, la méthode d'attribution par typage microbiologique, préalablement développée par le Danemark (Hald, Vose et al. 2004), a été adaptée dans 6 pays européens, dont la France. L'article qui suit présente ces démarches au travers des données de surveillance utilisées par chacun des pays, leurs caractéristiques (nombre de sources considérées, nombre de sérotypes inclus, mode de typage des souches...), les hypothèses posées et les résultats d'attribution obtenus. Il discute ensuite l'hétérogénéité des résultats d'attribution par source entre les différents pays, l'importance constante des cas exogènes (résultant d'un voyage) et l'absence de convergence pour certains pays.

Cet article est en préparation pour soumission à Foodborne Pathogens and Disease.

***Salmonella* source attribution in different European countries**

S.M. Pires^{1*}, G. Nichols², H. Whalstrom³, A. Kaesbohrer⁴, H. Spitznagel⁴, W. Van Pelt⁵, E. Snary⁶, P. Nally⁶, C. Frank⁷, J. David⁸, A. Baumann⁹, T. Hald¹

¹ National Food Institute, Technical University of Denmark, Lyngby, Denmark.

² Health Protection Agency Centre for Infections, 61 Colindale Avenue, London NW9 5EQ, United Kingdom

³ National Veterinary Institute, SE-75189 Uppsala, Sweden

⁴

⁵ Epidemiology and Surveillance Unit, National Institute for Public Health and the Environment, Bilthoven, The Netherlands.

⁶ Veterinary Laboratories Agency, Surrey, KT12 5JN, England.

⁷

⁸ French Food Safety Agency, BP 90203, 35302 Fougères, France

⁹ Department of Epidemiology, National Institute of Public Health, National Institute of Hygiene, Poland

* Corresponding author: Sara Monteiro Pires, National Food Institute, Technical University of Denmark. Mørkhøj Bygade 19, DK-2860 Søborg, Denmark. Tel: 45 5115 7464; Fax: +4572347028. Email: smpi@food.dtu.dk.

Abstract

Salmonella spp. is a major cause of human gastroenteritis in Europe. Several countries have implemented interventions to control *Salmonella* in food-producing animals and foods. However, knowledge on the impact of such interventions in the incidence of salmonellosis in humans is sparse. Under the scope of the "Human Illness Source Attribution" working group supported by Med Vet Net (WP 28), we adapted a previously developed microbial subtyping approach for the attribution of human salmonellosis to specific sources in six European countries. The model estimates the number of domestic and sporadic cases caused by different *Salmonella* subtypes as a function of the distribution of these subtypes in the animal-food sources, the amount of food source consumed, and the differences between subtypes and sources in their ability to cause disease. The approach was adapted to surveillance data from Germany, the United Kingdom, the Netherlands, Sweden, Poland, and France. Adaptations were made according to data features and availability, and the results were interpreted in light of the differences in public health, animal and food surveillance, and food consumption figures. National source attribution models were successfully developed for the UK, Sweden, the Netherlands and France; the attempted models for Germany and Poland did not provide appropriate convergence. Results suggest that the most important sources of human salmonellosis vary between countries, and that international travel is the single most important cause of disease in the majority of the investigated countries. A comparison of both methods and results shows that the successful adaptation of this approach is dependent on the quality and representativeness of the available data. The model is a useful tool to estimate the most important sources of salmonellosis in the countries, and we encourage other countries to apply source attribution approaches to assist risk management and public health decisions.

1 Introduction

Human salmonellosis is a major cause of disease in Europe. In 2006, the incidence of human salmonellosis in European Union Member States (MS) was 34.6 per 100,000 inhabitants (EFSA, 2007). A wide variability is observed amongst the countries, and incidence estimates ranged from 3.0 to 321.5 per 100,000 population in the different countries (EFSA, 2007). This variability also reflects differences in surveillance and reporting systems in European countries.

During the last decades, several countries have implemented intervention programmes to prevent and control foodborne diseases, particularly foodborne zoonoses (Wegener et al., 2003; EFSA, 2006). However, precise measure of the public health impact of such interventions has been difficult, in part because the contribution of specific sources to the burden of foodborne diseases is often unknown. To prioritize appropriate food safety interventions, it is crucial to attribute the human disease burden of each foodborne infection to specific sources (Pires et al., 2009)..

In order to get a better understanding of the mechanism behind the dynamics in the occurrence of *Salmonella* infections in humans, Hald *et al* (2004) developed a *microbial subtyping approach* for source attribution. The method estimates the number of human cases attributable to each of the major animal-food sources. The method is routinely applied by the Danish Zoonosis Centre, National Food Institute, and its results are published in the Annual Report on Zoonoses in Denmark (e.g. Anon., 2006). Microbial subtyping has proved to be a valuable tool in focusing food safety interventions to the appropriate animal reservoir in Denmark, and provides an example of potential synergy between quantitative risk assessment and public health surveillance (Wegener et al., 2003;Hald et al., 2004).

An attempt to adapt the model to other European countries' data was made under the scope of the "Human Illness Source Attribution" working group supported by Med Vet Net (WP 28). National source attribution models for *Salmonella* were developed for Germany, United Kingdom, the Netherlands, Sweden, Poland, and France. The method was adapted according to the availability and features of the data, and the interpretation of results took into account differences in public health processes, animal and food production, and food consumption figures. We present a comparison of the methods and results from these models and the model applied to Danish surveillance data.

2 Material and Methods

2.1 Model overview

The model is built on the basis of the method described by Hald *et al.* (2004), where *Salmonella* subtypes isolated from animals and foods in a given year are compared with subtypes isolated from humans. The objective is to estimate the number of human cases that can be attributed to each source. The method is developed under a Bayesian framework and allows for the estimation of the contribution of additional factors, other than the prevalence of each type on the food/ source and the amount of food consumed, to the number of people infected with a particular type. These factors account for the differences in the ability of different *Salmonella* subtypes to cause disease and of different sources to act as a vehicle for infection.

The basic model attribution equation is

$$\lambda_{ji} = P_{ij} * m_j * a_j * q_i \quad (\text{Equation 1}),$$

where λ_{ji} is the expected number of cases per type i and source j , P_{ij} is the prevalence of type i in source j , m_j is the amount of source available for consumption in the country, a_j is the source-dependent factor for source j , and q_i is the bacteria dependent factor for type i . The source and subtype related factors a_j and q_i were defined as uniform prior distributions. q_i for phage types within *S. Enteritidis* were fixed to 1, and all remaining q_i estimates are relative to this one.

2.2 Data used in the model

The original model attributes human sporadic and domestic *Salmonella* infections to specific sources based on the number of laboratory-confirmed infections caused by each *Salmonella* subtype, travel information for each case, prevalence of each subtype in the different sources, and amount of source available for consumption in the country. The number of outbreak cases per type (except 1 case) is subtracted from the total number of cases

reported per type to estimate the number of sporadic cases. The remaining outbreak-related cases are in the end added to the output of the model. Travel information and amount of food source available for consumption are not an absolute data requirement, and the model can be adapted without these parameters.

The variables and data differed between countries (Tables 1 and 2). This reflects differences in data availability with regards to typing of *Salmonella* isolates, travel information, number of sources tested, and testing of imported foods.

The sources included in the models also differed between the countries. These differences reflect not only availability of data, and thus coverage of national surveillance programmes, but also differences in consumption habits and in the importance of the sources. Cases were attributed to an "unknown" source if they were caused by subtypes not found in the sources considered in the model, or caused by subtypes grouped into "other serotypes", which the model could not attribute to any of the sources.

All models were run in the WinBUGS 1.4 software with five independent Markov chains of 40,000 iterations, each starting with a different set of initial values for the prior distributions. Convergence was monitored using the methods described by Gelman and Rubin (Gelman and Rubin, 1992) and was considered to have occurred when the variance between the different chains was not larger than the variance within each individual chain, and when the chains had reached a stable level.

The Swedish model

In the period between the 1st of July 2004 and the 31st of June 2006, 1,296 laboratory-confirmed sporadic cases, 227 outbreak-related cases and 5,859 travel-related cases were reported in Sweden. All *Salmonella* isolates were serotyped, and isolates of the serotypes Enteritidis and Typhimurium were phage typed. Salmonellosis sporadic cases were attributed to 10 sources: *imported food (including human foreign cases), pigs, cattle, layers, broilers, geese, passerine birds, seagulls, hedgehogs* and *unknown*. Data from travel related cases were included in the model as an addition to the available data on food. "Food" was considered as a single source, and grouped isolates from samples from different imported foods sampled at the retail level. The final model included 81 *Salmonella* subtypes, including "*Salmonella* others", which grouped subtypes that only caused few cases of disease and subtypes not found in any of the considered sources.

The UK model

In the United Kingdom, a total of 41,264 lab-confirmed cases were reported to the Health Protection Agency between 2003 and 2005. All *Salmonella* isolates were serotyped, and isolates of the serotypes Enteritidis and Typhimurium were phage typed. Human reported cases included information on travelling abroad, although there is evidence that this was not completed for all travel related cases. Human salmonellosis domestic sporadic cases were attributed to 6 domestic animals-food sources, *game bird* and *unknown*. The final model included 55 types (including "others").

The Dutch model

A total of 1,232 sporadic cases, reported in 2006, were included in the Dutch model and attributed to 4 sources: pigs, cattle, boilers and layers. All *Salmonella* isolates were serotyped, and isolates of the serotypes Enteritidis and Typhimurium were phage typed. The travel-related cases were extracted before modelling, and the proportion of the total salmonellosis cases attributed to *travelling* was added again to the output of the model. The dataset did not include human cases caused by subtypes not found in the animal sources. The final set included 48 types (including "others", which grouped the subtypes causing only few human cases).

The German model

44,294 human salmonellosis cases were reported in Germany. These were attributed to 6 sources: pork, beef, chicken, eggs, turkey and reptiles. All *Salmonella* isolates were serotyped, and isolates of the serotypes Enteritidis and Typhimurium were phage typed. Travel information was available for cases caused for all *Salmonella* serotypes, but not for phage types within *S. Enteritidis* and *S. Typhimurium*. Data on the amount of source available for consumption (m_j) of each source was included in the model; m_j for the source *reptiles* was assumed to be not available. The model included 67 types (including "others", which grouped isolates not found in the considered sources and infrequent subtypes).

The Polish model

The Polish dataset included 32,187 laboratory-confirmed cases, reported in 2005 and 2006. All isolates were serotyped, but no phage typing information was available. The data did not

include travel information. Nine sources, including domestic animal foods, aquatic products and wild animals, were considered in the model: pork, beef, poultry, eggs, shellfish, other animals, guinea fowl, buffalos and wild game.

The French model

The adaptation of the model to French data involved more substantial changes. The adaptation of the source account model to *Salmonella* French data was made at the French Food Safety Agency (AFSSA). The French dataset for 2005, includes 9076 domestic sporadic lab-confirmed human cases as well as active surveillance data on layers, broilers, turkeys, pigs and cattle and passive surveillance data on those 5 channels plus sheep, ducks, other poultry and sea products. Data features include the serotyping of all *Salmonella* isolates and further subtyping of *S. Enteritidis* and *S. Typhimurium* by clusters of antimicrobial resistance profiles. The model included travel information and consumption data for the included sources. Sporadic *Salmonella* infections are being attributed to pigs, cattle, broilers, layers, turkeys, ducks, sheep, other poultry and sea products. The human data are provided by the National Reference Center for Salmonella and the National Public Health Institute (Institut de Veille Sanitaire). Regarding the agro-food sector, the active data are provided by the Food Directorate (Agriculture Ministry) and the passive data, by the Salmonella Network (French Food Safety Agency).

The Danish model

For the purpose of comparison of methods and results, we present the output of the Danish model (as described above and by Hald et al. (2004)) applied to Danish *Salmonella* surveillance data from 2007. In this model, cases caused by 54 *Salmonella* subtypes (including "others", which grouped isolates not found in the considered sources and infrequent subtypes) were attributed to eight sources: pork, beef, table-eggs, chicken, imported pork, imported beef, imported chicken and imported turkey.

3 Results

Adaptations of the *Salmonella* source attribution model were successful for data from the United Kingdom, Sweden, the Netherlands, and France. The models from Germany and Poland did not provide appropriate convergence, and thus results were not analyzed and are not presented. *Salmonella* source attribution estimates for France are in progress (David et al., in preparation).

The proportion of cases attributed to the main sources varied considerably between the three countries (table 3). In the UK, the food-source causing a higher proportion of cases was chicken, followed by table eggs and pork. In contrast, the majority of the salmonellosis cases in the Netherlands were estimated to be attributed to table eggs, followed by beef and pork. In Sweden, the most important source was *food*, which comprises cases caused by imported foods; wildlife sources, and the animal sources layers, cattle and chicken, all causing a very low proportion of cases (<1%), followed in order of importance.

Results suggest that international travel was the single most important cause of salmonellosis in the UK and Sweden. Likewise, the majority of salmonellosis in Denmark was attributed to travel. The proportion of cases attributed to travel in Sweden and in the Netherlands correspond to reported travel-related cases, and were not estimated. The proportion of cases attributed to an unknown source varied between 20 and 37% in the UK, Sweden and Denmark. In the Netherlands, only around 4% of the cases were attributed to *unknown*, but these cases correspond only to infections caused by infrequent types and do not include cases caused by subtypes not isolated from the sources considered in the model. As a consequence, this estimate is not comparable with the previous ones. The proportion of outbreak-related cases was higher in Netherlands, representing around 11% of the total cases of reported salmonellosis.

The UK and the Swedish models attributed human salmonellosis to sources not included in the other models. In the UK, around 2% of the cases were attributed to lamb, and 1.4% to the consumption of turkey. The Swedish model attributed around 6.8% of human disease to imported foods, and cases were also attributed to wildlife sources; around 0.2% could be attributed to geese, 0.2% to pass birds, 0.09% to seagulls, and 0.3% to hedgehogs.

The relative importance of the source (a_j) and the subtype (q_i) related factors differed between the countries (table 4). Results are not directly comparable and should not be contrasted in absolute numbers. Chicken and eggs were on the top three of the most importance countries for two countries. Regarding the ability of *Salmonella* subtypes to

cause disease, the order of importance varied in all countries, and only *S. Newport* was common to the top subtypes with highest ability to cause disease in two of the countries (UK and the Netherlands).

4 Discussion

The Danish source attribution model was adapted to attribute human salmonellosis in the UK, Sweden and The Netherlands. The three models differed in the number of *Salmonella* subtypes, number of sources, and type of sources included. A comparison of results was only possible for the sources included in the three models. These estimates were further compared with the results of the original model, applied to attribute human salmonellosis in Denmark.

A comparison of the relative proportion of the cases attributed to the common sources included in the three models and in the Danish source attribution model revealed that the contribution of the main animal-food sources varied between countries. *Table-eggs* was estimated to be the most important source of *Salmonella* infections in the Netherlands and Denmark, and a major contributor for human salmonellosis in the UK; in Sweden, this source was estimated to be of minor importance. Chicken was estimated as the major source of salmonellosis in the UK, and results of the Dutch model suggest cattle is a very important source of disease in the country. Travel was estimated to be the single most important cause of disease in the UK, Sweden and Denmark; in the Netherlands, travel-related cases were not modelled, and cases reported with a history of travelling were added to the output of the model. The presented proportion could therefore be an underrepresentation of the true importance of international travel to human salmonellosis in the Netherlands.

In Sweden, the majority of the cases was acquired abroad or derived from the consumption of imported foods. A very low proportion of the cases was attributed to any of the food-animal sources. The low estimated proportion of cases attributed to domestic food-animals is explained by the very low prevalence of *Salmonella* in all food producing animals in the country. Sweden's *Salmonella* control program covering the entire food chain, from feed to fork, was implemented over 50 years ago, and interventions have proven successful in

reducing both the levels of *Salmonella* in the whole production chain and the burden of human salmonellosis attributed to domestic sources.

Even though the model provided appropriate convergence, the results of the Dutch model were not coherent with results from a previously developed source attribution approach that utilized the same data (Van Pelt et al., 1999; Valkenburgh et al., 2007). The principle of this method is very similar to the model developed by Hald et al. (2004), but the statistical methods differ. Despite these differences, we would expect results to be similar. However, the order of importance of the four sources included in the estimates differ, and the original Dutch approach attributes 29% of cases to table eggs, 21% to pigs, 11.6% to chicken and 11% to cattle compared to 41% for table eggs, 11% for pigs, 16% for cattle and 0.6% for chickens in this model.

The relative importance of the estimated source-dependent factors (a_j) was substantially variable in the three developed models and Denmark. In all models, these factors account for relative differences in food consumption, sensitivity of the surveillance programmes, variability of the sampling schemes and specific characteristics of the foods that may influence their ability to act as a vehicle for infection, e.g. fat content, preparation and processing, etc. It is also noteworthy that these estimates should not be compared in absolute numbers, especially between countries that do not include food consumption data in the models (e.g. UK, Sweden, the Netherlands) and countries that do (Denmark).

The *Salmonella* subtypes with highest ability to cause disease varied significantly in the three countries. These estimates account for differences in the survivability of the various *Salmonella* subtypes along the food chain, and potentially differences in their pathogenicity to humans. However, the method is a sort of "black box" model and the q_j values do not translate into virulence of the specific subtypes. In addition, it is expected the most important subtypes in different countries are a reflection of the occurrence of subtypes in those countries, as well of the different types of sources the population is exposed to.

The model developed for Poland did not converge. The available data had little discriminatory power (only serotype information), and merged information for multiple years. Despite the discriminatory power of the German data, which included serotyping and phage typing of the isolates, the model developed for Germany also did not provide appropriate convergence.

An adaptation of the source account model to *Salmonella* French data is in progress at the French Food Safety Agency.

Conclusion

The successful adaptation of the model is totally dependent on the quality and representativeness of the available data. Still, the approach has proved useful for the attribution of human salmonellosis, and therefore for the identification of the most important sources of *Salmonella* in different countries. We conclude that the presented source attribution approach is a valuable tool for the implementation of public health intervention strategies in different countries, and to point out areas for future research and data collection.

References

- Gelman, A., Rubin, D., 1992. Inference from iterative simulations using multiple sequences (with discussion). *Statistical Sciences* 7, 457-511.
- Hald, T., Vose, D., Wegener, H.C., Koupeev, T., 2004. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Anal.* 24, 255-269.
- Pires, S.M., Evers, E.G., van, P.W., Ayers, T., Scallan, E., Angulo, F.J., Havelaar, A., Hald, T., 2009. Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathog. Dis.* 6, 417-424.
- Valkenburgh, S., Oosterom, R., Van Stenvers, O., Aalten, M., Braks, M., Schimmer, B., Giessen, A., Van Pelt, W., Langelaar, M. Zoonoses and zoonotic agents in humans, food, animals and feed in the Netherlands 2003-2006. RIVM rapportnummer:330152001 ISBN-13, 978-90-6960-184-7. 2007.
Ref Type: Report
- Van Pelt, W., Van De Giessen, A.W., Leeuwen, W.J., Wannet, W., Henken, A.M., Evers, E.G. Oorsprong, omvang en kosten van humane salmonellose. Deel 1. Oorsprong van humane salmonellose met betrekking tot varken, rund, kip, ei en overige bronnen. *Infectieziekten Bulletin* , 240-243. 1999.
Ref Type: Report
- Wegener, H.C., Hald, T., Lo Fo, W.D., Madsen, M., Korsgaard, H., Bager, F., Gerner-Smidt, P., Molbak, K., 2003. Salmonella control programs in Denmark. *Emerg. Infect. Dis.* 9, 774-780.

Country	Typing		Travel	Food consumption data
	Serotyping	Phage typing		
UK	+	+	+	-
Sweden	+	+	NM*	-
Netherlands	+	+	NM*	-
France	+	+**	+	+
Germany	+	+	+	+
Poland	+	-	-	-
Denmark	+	+	+	+

*NM: not modelled

** : other subtyping method

+: Data included in the model

-: Data not included in the model

Table 1: Variables used in the different European models

Country	Sources				
	Food/ Dom. Animals	Food/ Game	Food/ Seafood	Food/ Imported	Wildlife
UK	+	+	-	-	-
Sweden	+	+	-	+	+
Netherlands	+	+	-	-	-
France	+	-	+	-	-
Germany	+	-	-	-	Reptiles
Poland	+	+	+	-	+
Denmark	+	-	-	+	-

*Imported food and “foreign cases” as a proxy for imported food

+: Data included in the model

-: Data not included in the model

Table 2: Sources included in the different European models

Source	UK	Sweden	Netherlands	Denmark
Pigs/pork	3.4 - 3.7	0.002 - 0.3	7.6 - 15.2	3.6 - 9.7
Cattle/Beef	0.1 - 0.25	0.003 - 0.3	12.3 - 20.1	0.2 - 1.6
Broilers/Chicken	15.2 - 15.8	0.008 - 0.2	0.2 - 1.8	0.1 - 1.8
Layers/Eggs	8.8 - 9.3	0.08 - 0.3	37.5 - 44.5	8.9 - 13.2
Turkey	1.1 - 1.7	NE	NE	NE
Lamb	2.0 - 2.3	NE	NE	NE
Game birds	0.6 - 0.7	NE	NE	NE
Food	NE	6.2 - 7.5	NE	NE
Wildlife†	NE	0.3-0.9	NE	NE
Travel	37.8	81.9	16.40	46.3
Outbreaks	5.8	2.9	10.80	4.4
Unknown	25.1 - 25.8	7.3 - 6.6	1.0 - 7.8	20.0 - 26.8

NE: not estimated

* "Unknown" in The Netherlands does not represent cases caused by types that were not identified in the described sources

** Travel related cases in the Netherlands were not modelled, but added to the final output of the model

†Includes geese, passerine birds, seagull and hedgehog

Table 3: Estimated proportion of cases of human salmonellosis attributed to specific sources in the UK, Sweden, the Netherlands and Denmark (95% CI)

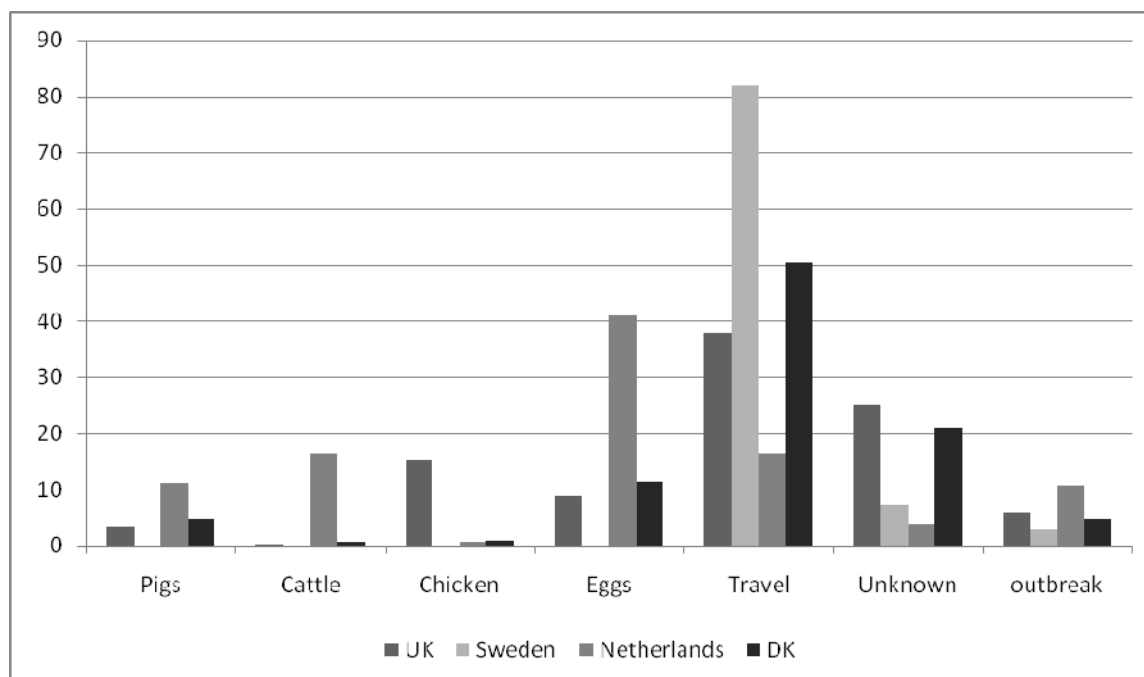


Fig. 1: Estimated proportion of cases attributed to specific sources in the UK, Sweden, The Netherlands and Denmark

* "Unknown" in The Netherlands does not represent cases caused by types that were not identified in the described sources

** Travel related cases in the Netherlands were not modelled, but added to the final output of the model

***Percentages do not add to 100%; some sources excluded

UK		a_j			q_i		
Chicken	15250	[14610,	15900]	S. Newport	2215	[1700, 2855]	
Eggs	10030	[9469,	10590]	S. Hadar	57.4	[34.1, 89.4]	
Lamb	3507	[2648,	4407]	S. Arizonae	44.7	[30.4, 63.3]	
Sweden				Sweden			
Chicken	26	[91.6,	0.3]	S. Napoli	214.9	[127, 332.2]	
Hedgehog	19.2	[88.0,	0.1]	S. Typhimurium			
				NST	130.2	[59.6, 231.5]	
Seagull	12.1	[80.9,	0.0]	S. Typhimurium			
				DT12	117.2	[57.7, 200.6]	
Netherlands				Netherlands			
Eggs	3.5	[3.2,	3.8]	S. Corvallis	162.2	[9.0, 468]	
Cattle	2.1	[1.4,	3.0]	S. Muenchen	38.4	[16.4, 75.3]	
Pig	0.3	[0.2,	0.4]	S. Newport	30.2	[12.0, 61.3]	
Denmark				Denmark			
				S. Typhimurium			
Beef	$4.21 \cdot 10^{-6}$	$[7.41 \cdot 10^{-9},$	$1.75 \cdot 10^{-5}]$	DT104 Q	243.3	[6.4, 852.5]	
Imported beef	$4.21 \cdot 10^{-6}$	$[7.41 \cdot 10^{-9},$	$1.75 \cdot 10^{-5}]$	S. Thompson	74.8	[22.1, 99.2]	
Eggs	$8.06 \cdot 10^{-7}$	$[6.47 \cdot 10^{-7},$	$9.73 \cdot 10^{-5}]$	S. Senftenberg	66.6	[22.0, 98.4]	

Table 4: Estimated source- (a_j) and food-dependent factors (q_i) for the UK, Sweden, the Netherlands and Denmark, mean and 95% CI

IV

Sensibilité du modèle aux priors informatifs

Quatrième article

Modèle bayésien d'attribution par typage microbiologique :
sensibilité à l'information *a priori* informative et une proposition

En finalisation pour soumission à Foodborne Pathogens and Disease

Le modèle a été utilisé avec succès au Danemark, en Grande-Bretagne et aux Pays-Bas. Cependant, les résultats obtenus pour ces pays sont apparus comme divergents quant aux sources principales, hors cas liés à un voyage. De plus, pour la Pologne et l'Allemagne, la convergence du modèle n'a pu être atteinte. La raison invoquée repose sur le manque de qualité et de représentativité des données, ce qui semble peu réaliste pour les données allemandes.

Le modèle proposé par les Danois est un modèle surparamétré et nécessite donc l'introduction d'information *a priori* « informative » sur certains paramètres. Cette information informative peut porter sur le paramètre source-dépendant (qui mesure la différence entre les sources quant à leur capacité à véhiculer les salmonelles), et/ou sur le paramètre type-dépendant (qui mesure la différence entre les types quant à leur capacité à causer des infections). L'objectif des travaux présentés ci-après était d'évaluer l'impact de l'information informative introduite dans le modèle.

Pour ce faire, les analyses ont été conduites à partir du jeu de données français, pour lequel seules 4 sources potentielles ont été considérées (poules pondeuses, poulets de chair, dindes et porcs). Différentes variantes du modèle ont été considérées : un modèle simple déterministe, qui sert de base de référence, le modèle originel assorti de quatre types d'information différents, ainsi que l'adaptation du modèle proposée par Mullner *et al* (2009)¹. Les quatre types d'information considérés correspondent à deux variantes de la paramétrisation utilisée dans la publication originelle (Hald, Vose et al. 2004) (le paramètre type-dépendant est fixé à une valeur arbitraire pour un type de référence), et deux variantes que nous proposons (les paramètres type-dépendants correspondants aux types spécifiques sont fixés soit à une valeur arbitraire soit à une valeur calculée à partir des données).

¹ Cf partie Objet et contexte de l'étude, § 5.2

Les résultats obtenus pour chacune de ces 6 variantes sont comparés en termes d'adéquation du modèle, de résultats d'attribution par source et de distribution *a posteriori* des paramètres source et type dépendants.

Le projet d'article suivant est prévu pour soumission à Foodborne Pathogens and Disease, après relecture par un traducteur.

Bayesian microbial subtyping attribution model: sensitivity to informative prior information and a proposition

J.M. David^{*1,2,3,4}, D. Guillemot^{2,3,4}, N. Bemrah⁵, A. Thébault⁵, C. Danan⁶, M. Chemaly⁷, FX. Weill⁸, N. Jourdan⁹, P. Sanders¹, L. Watier^{2,3,4}

1 French Food Safety Agency, BP 90203 Fougères, F-35302, France

2 INSERM, U 657, Paris, F-75015, France

3 Institut Pasteur, Pharmacoépidémiologie et Maladies Infectieuses, Paris, F-75015, France

4 Univ. Versailles Saint Quentin, Faculté de Médecine Paris Ile de France Ouest, F-78035, France

5 French Food Safety Agency, Maisons-Alfort, F-94701, France

6 French Food Safety Agency, Maisons-Alfort, F-94700, France

7 French Food Safety Agency, BP 53, Ploufragan, F-22440, France

8 Institut Pasteur, CNR des *Salmonella*, Laboratoire des Bactéries Pathogènes Entériques, Paris, F-75015, France

9 Institut de Veille Sanitaire, Saint-Maurice, F-94415, France

Corresponding author: Julie David, French Food Safety Agency, Institut Pasteur, BP 90203 35302 Fougères, France. Tel: +33 2 99 94 78 78; Fax: +33 2 99 94 78 99. Email: j.david@afssa.fr

Acknowledgements:

This study was supported by the Délégation Générale de l'Armement, Defence Ministry.

The authors wish to thank T. Hald and S.M. Pires for providing the code and for their support, all members of the MedVetNet WP28 Working Group for the useful discussions, all members of the CEB team at the French Food Safety Agency, Maisons Alfort, and of the *Salmonella* NRC at the Pasteur Institute for the subtyping analyses realized especially for this work and the members of the and of the Research Unit of Epidemiology and Welfare of Poultry and Rabbits at the French Food Safety Agency, Ploufragan. In addition, we wish to gratefully acknowledge Professor Lellouch for his constructive comments.

1 Introduction

Assessing the relative importance of the different reservoirs is a major issue regarding foodborne zoonoses such as salmonellosis or campylobacteriosis. As *Salmonella* is ubiquitous and present at a non negligible level in all the food-animal reservoirs and considering the fact that no fail-proof way to ensure the safety of food exists at any stage of the food chain (Allard 2002), controlling the presence of this pathogen from the farm level on seems critical (Sofos 2008). The attribution tool is thus a useful, let say indispensable tool to identify and prioritize interventions aiming at controlling it throughout the food chain (Batz, Doyle et al. 2005). Several methods are used worldwide to perform attribution: microbiological approaches (microbial subtyping approach and comparative exposure assessment), epidemiological approaches (case control studies), analyses of outbreak investigations as well as intervention studies and expert elicitation (Batz, Doyle et al. 2005; Pires, Evers et al. 2009). One of the most advanced models so far has been developed by Hald et al (Hald, Vose et al. 2004). This microbial subtyping approach tool allows attributing lab-confirmed human cases at the reservoir point to animal sources and is based on the serotyping and subtyping of *Salmonella* strains. Its principle is to compare the distribution of the types within the human cases and in the food sources, taking into account the consumption of each food source by the population. Its specificity is to take into account the differences between sources in their capacity to vehicle the pathogen and between the bacterial types in their capacity to induce infection, thanks to a bayesian framework with Markov Chain Monte Carlo simulations. However, as noticed by Mullner *et al*, the proposed model is not identifiable because of too many parameters to be estimated (Mullner, Jones et al. 2009). Thus, it is necessary to select *a priori* some parameters and fix them. In a Bayesian framework, including informative priors could have an important impact in the simulated joint posterior distribution of the parameters and thus on the marginal posterior distributions

of interest (Gilks, Richardson et al. 1996). As a consequence, informative information included in a model has to be in agreement with the literature or validated by a sensitivity analysis (Binkowitz and Wartenberg 2001). With the informative priors (selected parameters and corresponding values) proposed by Hald *et al*, some countries met difficulties when applying the model to meet convergence, despite the availability of apparently reliable data on the main animal channels and on the human cases (Pires, Nichols et al. 2008). To overcome such problems, Mullner *et al* proposed a modification consisting in a hierarchical modelling, allowing reduction in the number of parameters to estimate. However, this way of doing entails a large increase in the posterior credibility intervals due to weak information on the introduced hierarchical level.

In this work, we propose to study the impact of the chosen informative priors on the marginal posterior distributions of interest, for Hald's proposal and a data-based alternative. Results are also compared with Mullner's approach and with a proposed simple model. This work was applied to the 2005 French dataset.

2 Material and methods

2.1 Data

The microbial subtyping approach requires spatially and temporally related data on the distribution of *Salmonella* types in the human cases and the various food sources at the reservoir point (i.e. farm level or abattoir). Data on the consumption of the considered sources is also necessary.

The data on the human cases come from the *Salmonella* National Reference Centre (NRC) and the National Public Health Institute (InVS). The human cases included in the model are domestic sporadic cases registered in 2005. Known travellers, cases from the overseas territories and departments, and outbreak related cases are excluded on the basis of the

information given by NRC and InVS. The most recent year for which all the necessary information on the cases was available was 2005, it is thus the one considered as reference year in this work.

As to dispose of national representative prevalence data per *Salmonella* serotype in the animal sources, we used the data collected by the Food Directorate of the French Agriculture Ministry in the frame of the European baseline studies (layers, broilers, turkeys and pigs) and of the national surveillance plan of antimicrobial resistance in indicator and zoonotic bacteria in cattle (David, Danan et al. submitted). The data were collected in 2005 for layers and cattle, 2006 for broilers and 2007 for turkeys and pigs, based on a national representative sample of farms or carcasses. No equivalent national representative data were available in France in 2005 for broilers, turkeys and pigs. The strains were collected at the farm level for layers, broilers and turkeys and at the abattoir level for pigs and cattle. The prevalences were adjusted to flock size for layers, broilers and turkeys.

To optimize the attribution, Enteritidis and Typhimurium strains which represent more than 30% of the human cases each, have been further subtyped. The subtypes were defined through Multiple Correspondence Analysis (MCA) and mixed classification (Berge, Atwill et al. 2003) applied on the antimicrobial resistance profiles (disk diffusion method, amoxicillin, chloramphenicol, ceftazidim, gentamicin, kanamycin, nalidixic acid, streptomycin, sulfamides, sulfamides-trimethoprim, tetracyclin) of all Enteritidis and Typhimurium strains for the animal sources and a subset of strains for the human cases (92 strains out of 3536 (2.6%) for Typhimurium ; 102 out of 3138 (3.3%) for Enteritidis).

To be as close as possible to the effective consumption of the different sources by the French population, the consumption data used, come from the national individual survey on food consumption (INCA study) conducted by the French Food Safety Agency (AFSSA) in 1999 on 3 003 representative subjects above 3 years old (Volatier 2000). These results were

actualized on the basis of the data concerning the available amount of each source on the market published annually by the French Livestock Institute. This allows taking the evolution of the consumption from 1999 to 2005 into account.

2.2 Description of the models

2.2.1 Simple approach

A simple approach was used to estimate in a simple manner the number of human cases contaminated by a source. As mentioned by Zwietering, a simple model is well adapted to have an insight on the functionality of the model (Zwietering 2008) and thus, the attribution estimations obtained with the following simple model were used as reference.

This approach assumes that all the sources are equivalent vehicles for the pathogen and all the serotypes have the same capacity to induce infection. The expected number of human cases of type i ($i=1,\dots,I$) linked to a source j ($j=1,\dots,J$) (λ_{ij}) is proportional to the prevalence of type i in source j (p_{ij}) and to the exposition of the human population to source j , measured by the amount of source j consumed by the general population (M_j). The total observed number cases of type i (o_i) is distributed among the sources in which type i is present according to the relative weight:

$$\lambda_{ij} = \frac{p_{ij} \times M_j}{\sum_j p_{ij} \times M_j} \times o_i \quad (1)$$

The expected number of cases due to source j is obtained by summing the λ_{ij} on i .

As mentioned earlier, only a percentage of Enteritidis and Typhimurium cases are subtyped. It is thus necessary to introduce a reallocation step in the process. The types distribution observed within the cases that have been subtyped is used to allocate the cases with unknown subtype. This assumes similar subtype distribution for the cases subtyped and for

the cases not subtyped. The obtained reallocated numbers of cases per type are used in place of o_i . This way of doing is referred to as “deterministic reallocation”.

2.2.2 Bayesian approach

Hald model: With previous notations, the number of observed human cases (o_i) is assumed to be Poisson distributed:

$$o_i \sim \text{Poisson}(\sum_j \lambda_{ij}) \text{ with } \lambda_{ij} = M_j p_{ij} q_i a_j \quad (2)$$

Thus, the expected number of human cases due to a given type i in a given source j (λ_{ij}) depends on the prevalence of the type i in the source j (p_{ij}), as well as the consumption of the source j in the general population (M_j) as for the simple model, but two new parameters are introduced: a source dependant parameter (a_j) and a type dependant parameter (q_i). As defined by Hald *et al*, the type dependant factor (q_i) summarizes the characteristics of the serotype (survivability, virulence, pathogenicity, ...) which determine its capacity to cause an infection and the source dependant factor (a_j) summarizes the characteristics of the source (physical properties, preparation methods, processing procedures, ...) which determine its capacity to act as a vehicle for *Salmonella* (Hald, Vose et al. 2004).

The expected number of cases due to source j is obtained by summing the λ_{ij} on i .

From the writing of the model, an overparameterization appears, which concerns J parameters. Thus, informative priors have to be introduced for at least J parameters, which consists in first selecting the parameters to be fixed and then determining the constant values to which they are fixed. This step will concern the type-dependant parameters, which can be justified by the fact that there are only J source-parameters and moreover that the type-dependant parameters, have a stronger impact on the results than the source dependent parameters (Sarwari, Magder et al. 2001).

The untyped Typhimurium and Enteritidis cases were here reallocated in a bayesian way, assuming, as for the simple model, that the subtypes distributions are similar for subtyped

and not subtyped cases, and using Gamma distributions to reflect the uncertainty on the observed numbers of cases per type. For more details see Hald et al (2004).

Mullner model: To avoid including informative priors, the q_i s are modelled as random variables which follow a log normal distribution with a fixed constant mean and a precision parameter to be estimated (in the place of the q_i s). The gamma distribution was used for the precision parameter. From the simulated joint posterior distribution, marginal posterior distributions of the random variables q_i s can be recovered and were used to draw comparisons with the other models, even if their trace plots didn't appear to stabilize. Prior information used was the same as authors' proposal (Mullner, Jones et al. 2009).

Common features

Unlike Hald's work, where three sources of uncertainty concerned the human data (travel's proportion, outbreak related cases, Enteritidis and Typhimurium subtyping) which led to a complex model, our application exclusively contains one source of uncertainty (Enteritidis and Typhimurium subtyping). Accordingly, with the French dataset, the model results are easiest to interpret for what concerns the impact of the informative priors on the attribution results and on the parameters estimations (Zwietering 2008).

For the three considered approaches, it is assumed that cases with a history of travel or living in the overseas departments and territories have not acquired their Salmonellosis on the national metropolitan territory and, the way around, that domestic cases have been infected on the metropolitan territory. Moreover, we consider that the human cases infected by a type included in the model shall have been contaminated by one of the sources considered. The exposition to the sources is assumed to be similar in the general population and in the cases. And finally, these approaches require the microbiological types to be heterogeneously distributed among the sources considered.

Prior information

Informative priors

Two different configurations of the Hald model, each declined in two versions were studied. The first one corresponds to the original proposal fixing a reference type, the second focuses on specific types (i.e. types present in only one food-animal source).

In their work, Hald *et al* assumes that the q_i s are of equal value for the subtypes within the serotype Enteritidis (reference type) and within the serotype Typhimurium. For Enteritidis, the q_i s are moreover fixed to an arbitrary constant value while for Typhimurium a uniform prior distribution is chosen (Table 1). This proposal is referred to as "Reference-Type 1".

Enteritidis has been chosen by Hald *et al*, because it is the most frequent serotype within the human cases. Though, in the French dataset, the most frequent serotype is Typhimurium. Thus, another configuration has been studied, where Typhimurium and Enteritidis roles are inverted, Typhimurium becoming the reference type. This one is referred to as "Reference-Type 2". Both proposals assume that type dependant parameters for Enteritidis and for Typhimurium are independent of subtypes. Since Enteritidis and Typhimurium serotypes are both concerned by the reallocation process, an influence of the informative prior on this stage is to be expected.

We propose to fix specific types, not involved in the reallocation process as to prevent any possible interaction with this last one. For these serotypes, the link between the number of cases due to type i and its prevalence in the source is direct, which allows to determine a data-based value for q_i . However, as to study the impact of the way the constant values are defined, the type dependant parameters associated to specific types are either fixed to an arbitrary constant value (Specific-Types 1) or to data-based values (Specific-Types 2). These values were calculated as follow:

$$q_i = \frac{o_i}{\sum_i o_i} \times \frac{1}{p_{ij}} \quad (3)$$

where o_i is the observed number of cases due to type i , p_{ij} is the prevalence of type i in the unique source j . We thus used the percentage of human cases divided by the prevalence in the source as an indicator of the capacity of type i to cause an infection.

Non informative prior distributions for other parameters

Prior distributions for the a_j s and not fixed q s are uniform distributions with 0 as lower value. To assess that the upper value set allows encompassing all the possible values for the parameter, the posterior distributions are visually checked. If these distributions seem to be arbitrarily cut off, the uniform distributions are widened as indicated by Hald *et al.*

Constant values and non informative prior distribution's parameters

Arbitrary constant and prior distribution's parameters were chosen to optimize the reallocation process and the convergence of marginal posterior distributions of the a_j s et q s unknown parameters (table 1). They were determined as to assure first the compatibility of the reallocation with the observed proportions of subtypes and, second, to ensure that a minimum of the a_j s and q s have marginal posterior distribution arbitrarily cut off.

Adequate posterior results were easily obtained for both Specific-Types based configurations but they were partially and particularly hard to achieve for both Reference-Type based configurations, and higher values than those proposed by Hald were necessary. For Reference-Type 1, 5 unknown q parameters (14% of the q 's) and for reference-type 2, 7 unknown q parameters (19% of the q 's) have a marginal posterior distribution cut off. However, using widely dispersed starting values didn't modify the posterior marginal distributions of interest.

2.3 Software used

Analyses were performed using Excel[®] for the simple approach and Winbugs 1.4[®] for the bayesian models (Lunn, Thomas et al. 2000).

3 Results

3.1 Data

A total of 9 076 human cases were included in the dataset, as well as the prevalence results for 519 layer farms, 371 broiler farms, 331 turkey farms, 1 166 pig carcasses and 334 cattle carcasses (Figure 1). However, because of the very low prevalence of Salmonella in cattle (only 2.4%), inducing an important asymmetric density of the marginal posterior distribution with a small number of attributed cases (around 50), cattle was excluded as a source for this comparative study. The types are heterogeneously distributed among the sources according to Fisher exact tests.

Serotypes considered were those who were the most frequent among human cases (> 30 cases) and in sources (> 15% of the strains within a source) and those corresponding to specific types. The serotypes which didn't fulfil those criteria were included in "others" category. For Enteritidis and Typhimurium, 9 subtypes were defined each. But only 5 subtypes for Typhimurium and 3 for Enteritidis were included in the model. Those subtypes are the ones observed simultaneously in the human cases and at least in one animal source. The other subtypes were grouped in categories "other Typhimurium" and "other Enteritidis". Human cases belonging to the three "other" categories could not be considered for the attribution because, depending on the source, those categories have different types

compositions, which makes it impossible to distribute the relative cases according to prevalences that are not comparable. Thus, as these cases cannot be attributed to a source, 5 938 cases spread between 28 serotypes, 5 Typhimurium subtypes and 3 Enteritidis subtypes were considered for attribution and referred to as attributable cases.

Our studied sample contains two reference types including 8 subtypes (2 of which are specific), 12 specific types not included in the reference types and 14 other serotypes. Excluding Enteritidis and Typhimurium subtypes, a total of 5 types are specific to pigs, 2 to layers and 2 to broilers and 2 to turkeys (Table 2).

Finally, on the basis of the updated results of the INCA study, national consumption for 2005 was 82 301 tons for layers (eggs), 84 842 tons for broilers, 18 967 tons for turkeys and 161 971 tons for pigs.

3.2 Attribution

Results presented correspond to runs of 100 000 iterations of the Gibbs sampler with a thin of 25. Convergence diagnostics were satisfactory (Cowles and Carlin 1996; Brooks and Gelman 1998; Brooks and Roberts 1998b). From these runs, parameters estimates (posterior means, posterior variances and posterior 95% credibility intervals) were computed from the last 50 000 iterations.

We first performed comparisons corresponding to adequacy diagnosis. It concerned the number of cases per type and the total number of attributed cases, which were compared to the observed values. We then analyzed the expected numbers of cases per source with the simple model's results as reference. Finally, for models giving adequate results, the type and source dependant parameters were considered.

3.2.1 Adequacy of the model

Predicted and observed numbers of cases per type are in good agreement for Specific-Types 2 and for Mullner models, unlike both Reference-Type configurations and Specific-Types 1 (figure 2). Indeed, for both Reference-Type configurations, 5 subtypes are not well predicted. Regarding Typhimurium, subtypes 1, 2, 3 and 4 are concerned. The expected number of cases for Typhimurium globally is 113 for Reference-Type and 1 for Reference-Type 2, where 1 807 Typhimurium cases are observed. The discrepancy is less spectacular but still significant for Enteritidis, with 2 concerned subtypes, SE multiS and SE1. For Reference-Type 1, 1 796 and 869 cases are expected respectively for SE-multiS and SE1 vs 2 092 and 615 observed cases. For Reference-Type 2, 315 cases are expected for SE1 vs 615 observed. These weak fits are not related to the cut off of the marginal posteriors distributions of the q 's, except for SE1 in the Reference-Type 2 configuration.

For Specific-Types 1 configuration, almost all specific types are not well predicted (Stourbridge, Bovismorbificans, Oranienburg, Heidelberg, S 48:z4,z23:-, Havana, Ohio, Goldcoast, Bareilly, Aijobo). Though, in this last case, the convergence is good and the mismatch concerns only serotypes associated to a small number of cases (5% of attributable cases overall).

As a consequence, the total expected number of cases with Specific-Types 2 ($n = 5\,745$) and Mullner ($n = 6\,000$) models are in accordance with the 5 938 observed attributable domestic sporadic cases. This is also true for Specific-Types 1 ($n = 5\,745$) despite the discrepancy for some of the specific types. Finally, for the Reference-Type configurations, a decrease of around 30% in attributed cases is observed (4 122 expected cases for Reference-Type 1 and 4 079, for Reference-Type 2).

3.2.2 Number of cases per source

According to the model, posterior attribution means per food source can be very different (figure 3). Only two proposals are in accordance with the proposed simple model, that is Specific-Type 2 and Mullner's models. In those cases, considering marginal posterior means or percentages, ranking are similar for the most important sources: layers as the main source for human cases and pigs at the second place (Tableau 3). The ranking for turkeys and broilers is not so clear. However, in both models these sources are not significantly different. Though, with Mullner's model, no firm conclusion on the relative importance of the sources can be made, because of very wide 95% CI (95% Credibility Intervals).

For the other models, the results are quite heterogeneous since presenting contrasting marginal posterior means, some of which seem unlikely. For example, when applying Reference-Type 1, less than 10% of the cases should be attributed to layers which is not compatible with literature (Mølbak and Neimann 2002; Hald, Vose et al. 2004; Hennessy, Cheng et al. 2004; Greig and Ravel 2009), and with the high proportion of Enteritidis strains observed in this source (Figure 1).

3.2.3 Type and source dependant parameters

The results for those parameters are presented only for Specific-Types 2 and Mullner models which have a good adequacy and attribution results in accordance with the simple model.

The ranking and even the posterior means of all type-dependant parameters are quite similar for both models, except for subtype ST2 (figure 4). Though, the 95% CI are so wide for this subtype that the difference of the posterior means for the two models isn't significant. In fact, wide credibility intervals related to the reallocation process are observed for all the subtypes within Enteritidis and to a less extent within Typhimurium.

The global agreement of q values between both models and especially for the specific types, strengthens the appropriateness of the proposed data-based informative values of the type dependant parameters for the specific serotypes. What's more, we compared the posterior means obtained with the Specific-Types 2 model for the 2 specific types not fixed because involved in the reallocation (ST2 and SE2), with the constant value that would have been calculated as defined in equation (3) (for the subtypes, the proportion of observed cases is obtained using the deterministic reallocation). The calculated value and posterior means were respectively 19.78 versus 14.01 (95%CI 0.68-37.17) for ST2 and 4.96 versus 4.30 (95%CI 0.17-12.76) for SE2, which shows a good agreement between the calculated and estimated values.

Several serotypes appear to have high posterior means, or calculated values. The Enteritidis subtypes, especially SE-multiS and SE1, ST2 and some of the specific types, i.e. Stourbridge (Pigs-related), Bovismorbificans (Pigs-related), Oranienburg (Layers-related), Heidelberg (Broilers-related) and Havana (Layers-related).

The source dependant parameters present, for both models, highest posterior means for broilers (Mullner: 1.11 [0.21 – 3.49], Reference-Types 2: 0.41 [0.30 – 0.55]). As the 95% CI for Mullner is very wide, the source-dependant parameter for broilers is not significantly different between both models. For what concerns the other sources. They are smallest than the broilers-dependant factors and near from each other (Mullner: 0.08 to 0.11, Specific-Types 2: 0.05 to 0.09). Within Specific-Types 2 model, the relative a factors are significantly smallest than the broilers-dependant one. Thus no firm conclusion can be drawn from the Mullner estimations, but with the Specific-Types 2 configuration, the source-dependant factor relative to broilers appears significantly higher than the three others.

4 Discussion

The dataset gathered for this work suffers from some insufficiencies. The human cases are collected on a passive way and on a voluntary basis, as sporadic Salmonellosis isn't submitted to mandatory declaration. Though, the coverage of the NRC database is estimated to 40% of the confirmed cases and its representativeness has been estimated as good (David, Danan et al. submitted). It thus constitutes a reasonable basis for such a study.

The data gathered for the sources are not timely consistent with the human ones, but they are representative and give access to prevalences per serotype and subtype, which wasn't possible with other sources of data on food-animals. As the aim of this work was to understand the functionality of the model and not to produce attribution estimates for France, we still used them. Though, the estimations provided by the models must be considered with caution in view of the quality of the dataset, the exclusion of cattle as a source and the underlying assumptions.

The overparameterization of Hald model has been originally treated by assuming equality in some type dependant parameters, while fixing a reference serotype to an arbitrary constant value. Recently, Mullner proposed to reduce the number of parameters to be estimated by introducing a hierarchical structure for the type dependant parameters despite of their low number. Our proposal targets the informative prior as did Hald *et al*, but is based on specific types rather than on a chosen reference type. We use the fact that, as specific types are present in a unique source, the link between the prevalence and the proportion of the human cases induced is direct. The values, to which they are fixed, are data-based and shall reflect their capacity to induce infection measured in percentage of cases per prevalence rate. These values are in agreement with marginal posterior means of Mullner's model which

comforts our proposal. For simplicity, all specific types were fixed, but, fixing only J specific types among which at least one in each source, led to similar results.

As expected, the bayesian attribution model appears very sensitive to informative priors and the way the priors are defined. Even when complying with Hald's proposal, that is fixing Enteritidis (Reference Type 1) or Typhimurium (Reference Type 2) which are almost equally frequent within the human cases, the attribution results are spectacularly different.

The Specific-Types based information has thus several advantages. Namely, in regard to the Reference-Type solution which fixes the most frequent serotypes (over 30% of the cases in our dataset), when fixing specific serotypes, only a small proportion of the human cases (5% in our dataset) is concerned. With no certainty on the accuracy of the constant values chosen to fix the relative q parameters, the potential direct influence is thus minimized. This also allows avoiding any interaction with the reallocation process in which the reference types are implied. Namely, when fixing a reference type dependant parameter, the reallocation process was shown to be dependant of the combination: constant value / upper value of the uniform distributions of (q_i) . Moreover, when fixing parameters to arbitrary values, the repartition of the cases among the sources isn't realistic, the total expected number of cases isn't in accordance with the number of attributable cases and the expected numbers of cases per type aren't in agreement with the observations for the fixed types, whichever those are (reference types or specific types). On the contrary, introducing data-based informative priors on specific types not involved in the reallocation seems a good alternative. Another point is that the Specific-Types based solution doesn't require to make any assumption on the equality of the type-dependant parameters for the subtypes within Enteritidis and Typhimurium. As we used the antimicrobial profiles as subtyping tool and as some resistance traits have been shown to be linked to virulence factors especially in Typhimurium (Martinez and Baquero 2002; Mølbak 2004; Foley and Lynne 2008), this is of

primarily importance. Regarding Mullner approach, the obtained results are comparable, but the convergence is easier to meet with the Specific-Types based solution and the credibility intervals are much nearer?, allowing concluding on the relative importance of the sources, which wasn't possible with Mullner model for our dataset.

Though, this approach requires disposing of as many specific types other than Enteritidis and Typhimurium as food-sources included, which for example was not the case in the Danish dataset (Hald, Vose et al. 2004), but maybe because specific serotypes were among the less frequent types categorized as "others". If no such specific serotypes can be found even in the less frequent ones, then a solution would be to apply the simple model, except that including the source-dependant and the type-dependant parameters is a key point to enhance the source attribution estimates, as differences between the sources in their capacity to vehicle the pathogen and the differences between types to induce infection are described in the literature (Blaser and Newman 1982; D'Aoust 1989; Sarwari, Magder et al. 2001; Coleman, Marks et al. 2004; Bollaerts, Aerts et al. 2008; Jones, Ingram et al. 2008). Another limit of this approach was common with Hald's and Mullner's proposals and linked to the dataset. The subset of Enteritidis and Typhimurium strains tested for antimicrobial resistance in the human dataset represents less than 4% of the total number of Enteritidis and Typhimurium cases. This leads to wide credibility intervals for the posterior distribution of the relative type dependant parameters and for the predicted numbers of cases per type and source, whichever the model was. Testing more strains for antimicrobial resistance could greatly enhance the predictions of all models.

As a conclusion, we proposed alternative informative priors, based on specific types, to be used in the Hald model. For datasets comprising specific types not involved in a reallocation process, our proposal allows enhancing the convergence and avoiding biases in the results of

the model. The obtained results are consistent with the simple model and allow drawing significant differences between the sources. We thus would recommend using the Specific-Types based informative prior when possible.

However, this approach could be enhanced in several ways. First of all, the bayesian framework allows to introduce uncertainty on the parameters. Thus, low flat distributions centred on the proposed data-based values could be introduced to replace the constant values to which the parameters are fixed. A further step in this approach would be to find exogenous information reflecting the infectious capacity of the concerned types (dose-response relationship, virulence or pathogenicity) to define those values.

Another path to enhance the quality of the model's predictions would be to use proportionality parameters in the informative priors set on the specific-types dependant parameters associated to the same source. At last, when running the model, the source and the type dependant parameters appeared inter-dependant, which is consistent with the known specificity of the dose-illness relationship for a serotype-food matrix combination (Bollaerts, Aerts et al. 2008). It thus would be interesting to initiate a reflection on how to introduce interactions between both parameters in the model.

Finally, a recent paper emphasized the possibility to obtain convergent results with inappropriate priors without warning from the BUGS software (Lunn, Spiegelhalter et al. 2009). Thus, it appeared in this work that the sensitivity analysis was essential to judge of the appropriateness of the approaches for a given dataset and that some apparently logical assumptions led to inappropriate results.

References

- Allard, D. G. (2002). "The 'farm to plate' approach to food safety - Everyone's business." *Canadian Journal of Infectious Diseases* 13(3): 185-190.
- Batz, M. B., M. P. Doyle, et al. (2005). "Attributing illness to food." *Emerging Infectious Diseases* 11(7): 993-999.
- Berge, A. C. B., E. R. Atwill, et al. (2003). "Assessing antibiotic resistance in fecal *Escherichia coli* in young calves using cluster analysis techniques." *Preventive Veterinary Medicine* 61(2): 91-102.
- Binkowitz, B. S. and D. Wartenberg (2001). "Disparity in quantitative risk assessment: a review of input distributions." *Risk Anal* 21(1): 75-90.
- Blaser, M. J. and L. S. Newman (1982). "A review of human salmonellosis: I. Infective dose." *Reviews of Infectious Diseases* 4(6): 1096-1106.
- Bollaerts, K., M. Aerts, et al. (2008). "Human salmonellosis: estimation of dose-illness from outbreak data." *Risk Anal* 28(2): 427-40.
- Brooks, S. P. and A. Gelman (1998). "General methods for monitoring convergence of iterative simulations." *Journal of Computational and Graphical Statistics* 7(4): 434-455.
- Brooks, S. P. and G. O. Roberts (1998b). "Convergence assessment techniques for Markov chain Monte Carlo." *Statistics and Computing* 8(4): 319-335.
- Coleman, M. E., H. M. Marks, et al. (2004). "Discerning strain effects in microbial dose-response data." *J Toxicol Environ Health A* 67(8-10): 667-85.
- Cowles, M. K. and B. P. Carlin (1996). "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review." *Journal of the American Statistical Association* 91(434): 883-904.
- D'Aoust, J. (1989). *Salmonella. Foodborne Bacterial Pathogens*. M. P. Doyle. New York, Marcel Dekker: 327-445.
- David, J. M., C. Danan, et al. (submitted). "Structure of the French farm to table surveillance system for *Salmonella*."
- Foley, S. L. and A. M. Lynne (2008). "Food animal-associated *Salmonella* challenges: pathogenicity and antimicrobial resistance." *Journal of animal science* 86(14 Suppl).
- Gilks, W. R., S. Richardson, et al. (1996). *Markov Chain Monte Carlo in practice: Interdisciplinary Statistics*. London, Chapman & Hall/CRC.
- Greig, J. D. and A. Ravel (2009). "Analysis of foodborne outbreak data reported internationally for source attribution." *International Journal of Food Microbiology* 130(2): 77-87.

- Hald, T., D. Vose, et al. (2004). "A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis." *Risk Anal* 24(1): 255-69.
- Hennessy, T. W., L. H. Cheng, et al. (2004). "Egg consumption is the principal risk factor for sporadic *Salmonella* serotype heidelberg infections: A case-control study in foodnet sites." *Clinical Infectious Diseases* 38(SUPPL. 3).
- Jones, T. F., L. A. Ingram, et al. (2008). "Salmonellosis outcomes differ substantially by serotype." *J Infect Dis* 198(1): 109-14.
- Lunn, D., D. Spiegelhalter, et al. (2009). "The BUGS project: Evolution, critique and future directions." *Stat Med.*(28): 3049-3067.
- Lunn, D. J., A. Thomas, et al. (2000). "WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility." *Statistics and Computing* 10(4): 325-337.
- Martinez, J. L. and F. Baquero (2002). "Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance." *Clin Microbiol Rev* 15(4): 647-79.
- Mølbak, K. (2004). "Spread of resistant bacteria and resistance genes from animals to humans - The public health consequences." *Journal of Veterinary Medicine Series B: Infectious Diseases and Veterinary Public Health* 51(8-9): 364-369.
- Mølbak, K. and J. Neimann (2002). "Risk factors for sporadic infection with salmonella enteritidis, Denmark, 1997-1999." *American Journal of Epidemiology* 156(7): 654-661.
- Mullner, P., G. Jones, et al. (2009). "Source attribution of food-borne zoonoses in New Zealand: A modified hald model." *Risk Analysis* 29(7): 970-984.
- Pires, S. M., E. G. Evers, et al. (2009). "Attributing the human disease burden of foodborne infections to specific sources." *Foodborne Pathog Dis* 6(4): 417-24.
- Pires, S. M., G. Nichols, et al. (2008). *Salmonella* source attribution in different European countries. 21st International ICFMH Symposium: "Evolving Microbial Food Quality and Safety". Aberdeen.
- Sarwari, A. R., L. S. Magder, et al. (2001). "Serotype distribution of *Salmonella* isolates from food animals after slaughter differs from that of isolates found in humans." *Journal of Infectious Diseases* 183(8): 1295-1299.
- Sofos, J. N. (2008). "Challenges to meat safety in the 21st century." *Meat Science* 78(1-2): 3-13.
- Volatier, J. L. (2000). Enquête nationale sur les consommations alimentaires. Tec&Doc, CREDOC-AFSSA-DGAL.
- Zwietering, M. H. (2008). "Quantitative risk assessment: Is more complex always better? Simple is not stupid and complex is not always more correct." *Int J Food Microbiol*.

Model name	Prior distribution for a	Informative and non informative distribution for q*	
		Types	Prior
Reference Type 1	Uniform distribution: U(0,10)	Enteritidis	Constant value : 10
		Typhimurium	Subtypes equals Uniform distribution: U(0,1000)
		Others	Uniform distribution: U(0,1000)
Reference Type 2	Uniform distribution: U(0,10)	Typhimurium	Constant value: 10
		Enteritidis	Subtypes equals Uniform distribution: U(0,10 000)
		Others	Uniform distribution: U(0,10 000)
Specific Types 1	Uniform distribution: U(0,100)	Specific types other than Enteritidis and Typhimurium	Constant value: 1
		Others	Uniform distribution: U(0,100)
Specific Types 2	Uniform distribution: U(0,100)	Specific types other than Enteritidis and Typhimurium	Specific value : $q_i = \frac{o_i}{\sum_i o_i} \times \frac{1}{p_{ij}}$
		Others	Uniform distribution: U(0,100)

Table 1: Informative and non informative prior information for Hald's model bold indicates informative information.

Type	Source	p	Number of cases	Specific value**
Heidelberg	Broilers	0.02%	30	19.240
Ohio	Broilers	0.07%	12	1.889
SE2	Layers	0.20%	92*	-
Oranienburg	Layers	0.09%	31	3.892
Havana	Layers	0.04%	16	4.407
Bareilly	Layers	2.95%	8	0.030
ST2	Pigs	0.09%	154*	-
Brandenburg	Pigs	0.34%	71	2.280
Stourbridge	Pigs	0.09%	42	5.396
Bovismorbificans	Pigs	0.09%	40	5.139
S 48:z4,z23:-	Pigs	0.18%	24	1.469
Goldcoast	Pigs	0.09%	9	1.102
Muenster	Turkeys	0.38%	7	0.201
Aijobo	Turkeys	0.04%	2	0.551

Table 2: Specific types characteristics

*reallocated number of cases for the subtypes, using the deterministic reallocation

** definition in table 1

p : prevalence

Source	Simple model	Specific-Types 2		Mullner	
	Percentage	Posterior percentage	95% CI	Posterior percentage	95% CI
Layers	53.3	53.5	46.0 - 60.0	40.0	17.6 - 60.0
Pigs	32.2	25.8	20.7 - 31.1	34.3	21.5 - 45.9
Broilers	8.7	8.0	5.6 - 9.5	13.6	4.7 - 27.5
Turkeys	5.8	12.7	6.9 - 19.2	12.1	2.0 - 45.9

Table 3: Posterior percentages of expected cases per source for the Specific-Types 2 and Mullner models in comparison to the simple approach

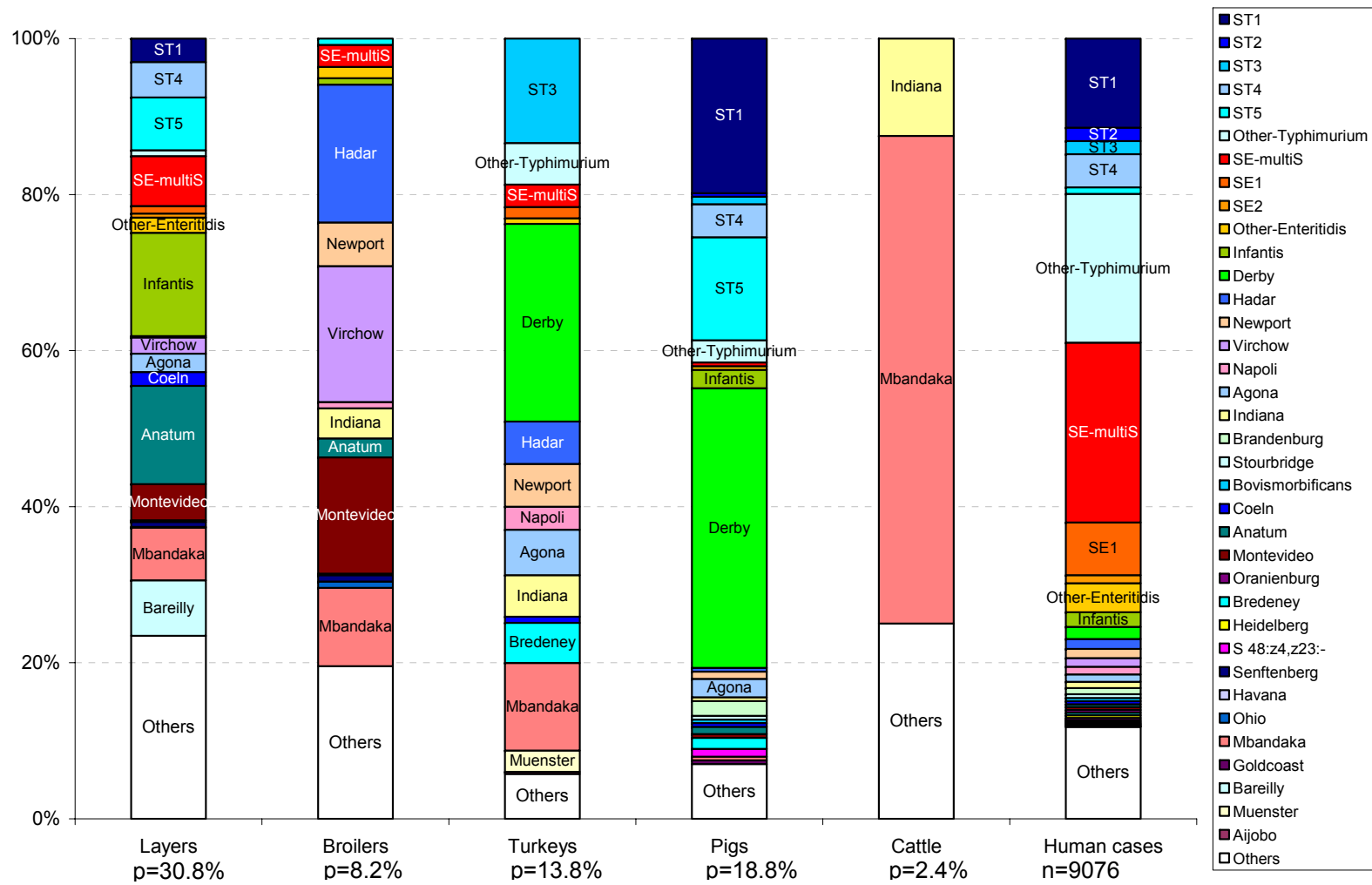


Figure 1: Surveillance results, types repartition in the sources and among the human cases, p indicated prevalence of the sources, n indicated the total number of human cases.

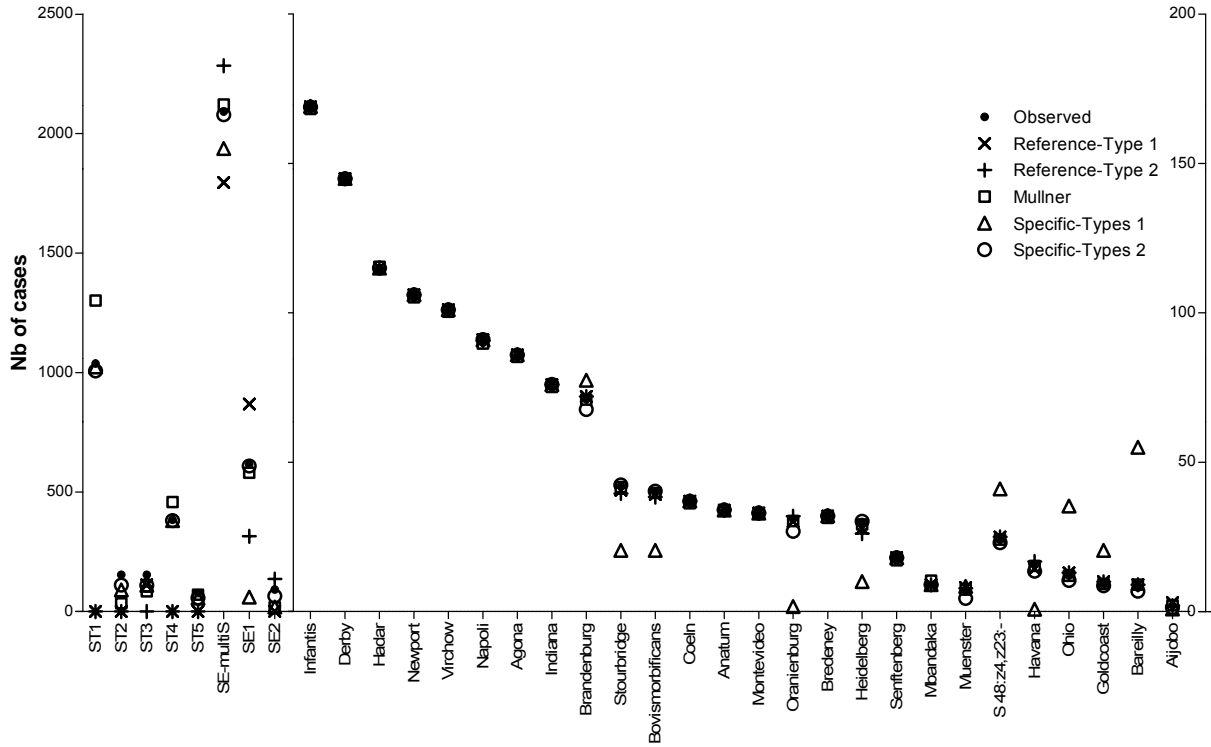


Figure 2: Comparison of posterior predicted cases per type and observed ones
 The observed numbers of cases for the subtypes are those obtained with the deterministic reallocation

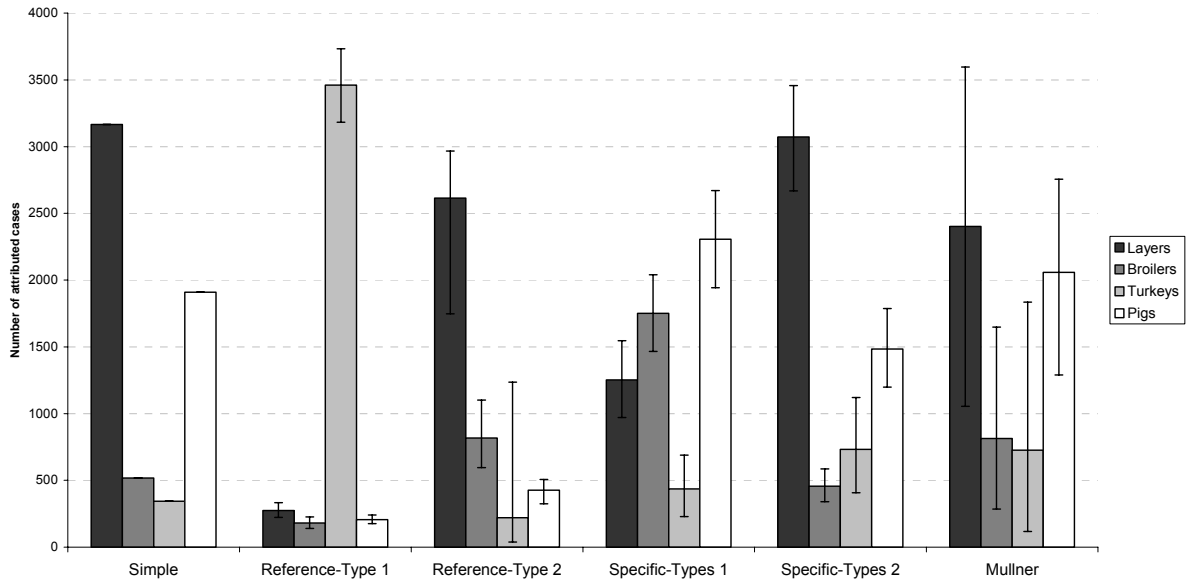


Figure 3: Attribution results for the simple model, the four parameterizations and Mullner's model

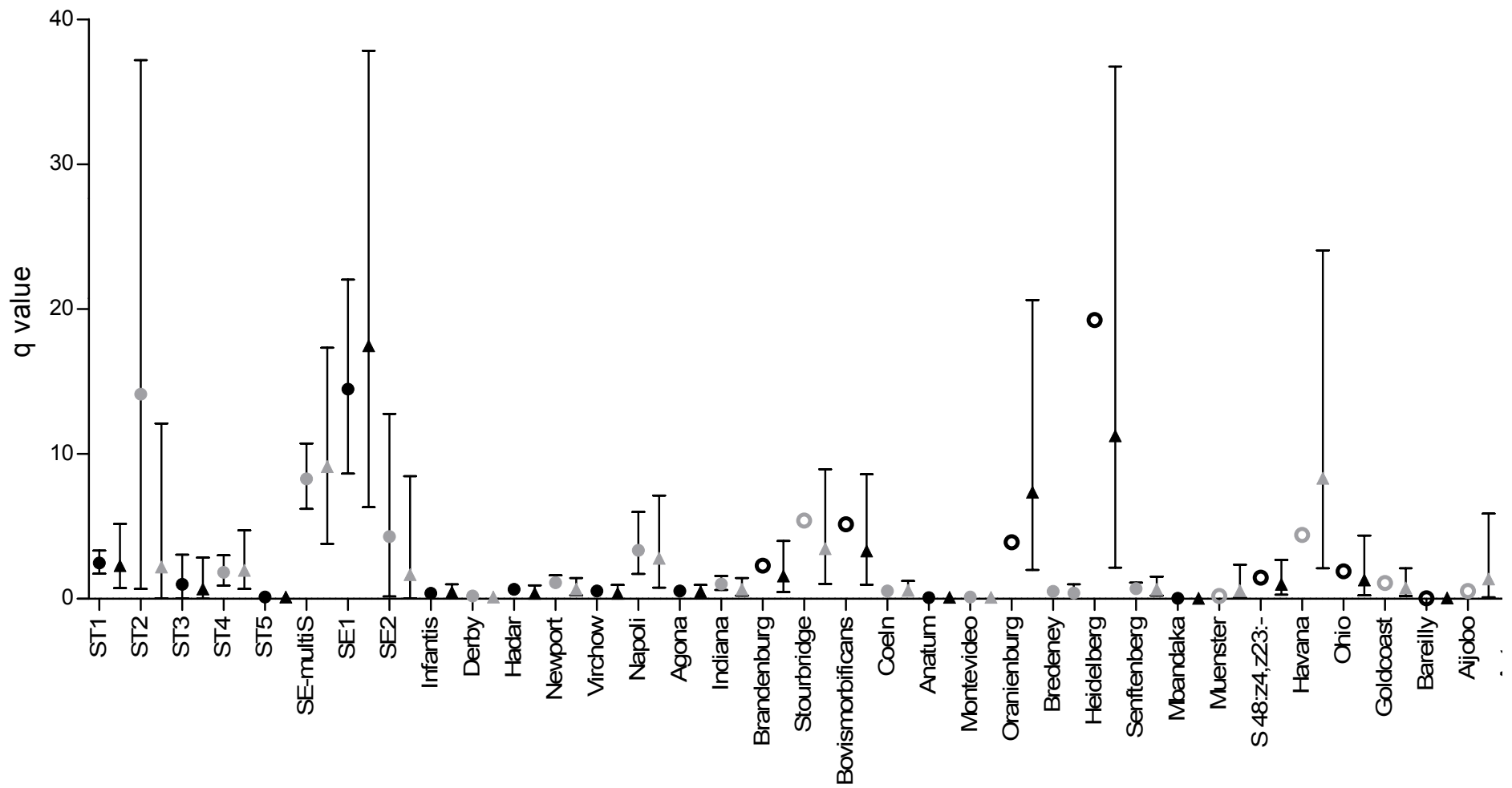


Figure 4: Posterior means and 95% CI of bacteria dependant parameters
 Circles are for Specific-types 2 posterior means; empty circles indicate the fixed values; Triangles are set for Mullner's model posterior means.

V

Impact de la qualité des données sur les résultats obtenus avec le modèle

Cinquième article

Attribution des cas humains français de salmonellose aux principales sources alimentaires selon le type de données de surveillance

Article en préparation pour soumission à Foodborne Pathogens and Disease

De façon attendue, l'information a priori informative a un impact non négligeable sur les résultats d'attribution, en particulier lorsque cette information est choisie de façon arbitraire. Parmi les paramétrisations évaluées, la proposition basée sur les types microbiologiques spécifiques à certaines espèces animales a permis d'obtenir des résultats convergents, et cohérents avec la méthode déterministe et les facteurs de risques connus dans la littérature. De plus, les résultats produits quant à l'importance relative des sources, sont significatifs. Cette approche a donc été utilisée pour évaluer l'impact de la qualité des données sur les résultats d'attribution.

En effet, comme présenté au chapitre II, le système de surveillance français permet de disposer parallèlement de données de surveillance active et de données de surveillance passive concernant la contamination des sources animales. La différence entre ces deux types de données concerne (i) la représentativité, structurelle pour les données de surveillance active, inconnue pour les données de surveillance passive, (ii) l'indicateur, prévalences pour les données actives, proportions pour les données de surveillance passive, et enfin (iii) le nombre de sources prises en comptes, 5 pour le jeu de données actives, 9 pour le jeu de données passives. L'impact des caractéristiques (ii) et (iii) est évalué individuellement à l'aide du modèle simple précédemment utilisé (Cf chapitre IV). Enfin, les résultats obtenus avec le modèle bayésien ont été comparés pour le jeu de données comprenant les données de surveillance active d'une part et, celui contenant les données de surveillance passive d'autre part.

Le projet d'article présenté est en cours de rédaction pour soumission à Foodborne Pathogens and Disease. Les références à l'article précédent sont notées de la façon suivante : (David, informative information).

Attribution of the French human Salmonellosis cases to the main food-sources according to the type of surveillance data

J.M. David^{*1,2,3,4}, D. Guillemot^{2,3,4}, N. Bemrah⁵, A. Brisabois⁶, M. Denis⁷, FX. Weill⁸, N. Jourdan⁹, P. Sanders¹, L. Watier^{2,3,4}

1 French Food Safety Agency, BP 90203 Fougères, F-35302, France

2 INSERM, U 657, Paris, F-75015, France

3 Institut Pasteur, Pharmacoépidémiologie et Maladies Infectieuses, Paris, F-75015, France

4 Univ. Versailles Saint Quentin, Faculté de Médecine Paris Ile de France Ouest, F-78035, France

5 French Food Safety Agency, Maisons-Alfort, F-94701, France

6 French Food Safety Agency, Maisons-Alfort, F-94700, France

7 French Food Safety Agency, BP 53, Ploufragan, F-22440, France

8 Institut Pasteur, CNR des *Salmonella*, Laboratoire des Bactéries Pathogènes Entériques, Paris, F-75015, France

9 Institut de Veille Sanitaire, Saint-Maurice, F-94415, France

Corresponding author: Julie David, French Food Safety Agency, Institut Pasteur, BP 90203 35302 Fougères, France. Tel: +33 2 99 94 78 78; Fax: +33 2 99 94 78 99. Email: j.david@afssa.fr

Acknowledgements:

This study was supported by the Délégation Générale de l'Armement, Defence Ministry.

The authors wish to thank T. Hald and S.M. Pires for providing the code and for their support, all members of the MedVetNet WP28 Working Group for the useful discussions, all members of the CEB team at the French Food Safety Agency, Maisons Alfort, and of the *Salmonella* NRC at the Pasteur Institute for the subtyping analyses realized especially for this work and the members of the and of the Research Unit of Epidemiology and Welfare of Poultry and Rabbits at the French Food Safety Agency, Ploufragan. In addition, we wish to gratefully acknowledge Professor Lellouch for his constructive comments.

1 Introduction

Salmonella is the first recorded bacterial cause of foodborne gastroenteritis in France, as well as the first cause of relative hospitalizations and deaths (InVS 2004). It is a major public health concern and as such gave rise to a nationwide intervention at the farm level targeting Enteritidis and Typhimurium serotypes in the breeding flocks of laying hens and broilers from 1999 on. The *Gallus gallus* plan has been successful at mitigating the number of human cases (Poirier, Watier et al. 2008), as other similar plans conducted in other European countries (Edel 1994; Wegener, Hald et al. 2003), showing the efficiency of intervening at an early point in the food chain. Though, the challenge remains as *Salmonella* is still a first plan public health issue (Sofos 2008). Attributing human cases to the potential food-sources is a key step to identify, prioritize and assess the impact of interventions from the farm level on (Batz, Doyle et al. 2005). The microbial subtyping attribution approach, which allows attributing human cases at the reservoir point, has already been applied successfully for such purposes in Denmark. It relies on the model published in 2004 by Hald *et al*, and which requires representative data on microbial type repartition both in human and in food animals (Hald, Vose et al. 2004; Mullner, Jones et al. 2009).

To answer such issues, the French integrated surveillance system for *Salmonella* has been developed in the European and national regulatory frame (David, Danan et al. submitted). It covers the whole food-chain from the breeding farms to the human cases and is thus well adapted to the microbial subtyping attribution approach.

Many actors are involved in this surveillance system, which allows disposing of different and complementary expertises (David, Danan et al. submitted). The data on the human cases are collected by the National Reference Centre (NRC), and the National Public Health Institute (InVS) centralizes the data on outbreak related cases which are submitted to mandatory declaration. For what concerns the data on the sources contamination, several databases are available. For this work, we selected the data collected by the Food Directory

of the French Agriculture Ministry (DGAL) in the frame of the institutional surveillance. We also used data collected in the frame of a generalist passive surveillance network (*Salmonella* Network, SN) of the French Food Safety Agency (AFSSA). The first dataset gathers representative prevalence data per microbiological type for the main channels, ideal for attribution. The passive dataset of SN, allows disposing of data on supplementary food-animal sources, which can be important as one of the assumptions of the attribution model is that all cases are linked exclusively to the included sources (Pires, Evers et al. 2009). We thus applied the microbial subtyping approach with both types of data, as to assess the opportunity to use the passive data in such an approach as a complement for active data, and studied the impact of the differences between the two kind of animal data on the model results. For this purpose, two different datasets were constituted, gathering the human data on Salmonellosis cases and either active surveillance data or passive surveillance data on the food-animal sources.

The impact of some of the data characteristics on the attribution results has first been studied, using a simple model (David, informative information), which allows having a more immediate insight on the interaction between the data and the model (Zwietering 2008). The general comparison of the attribution results obtained according to the surveillance type, was made using the Bayesian model described by Hald *et al* (2004), with an adapted informative prior, as proposed in a former paper (David, informative information).

2 Materials and methods

2.1 Data

2.1.1 Human cases

The human cases data were retrieved from NRC and InVS databases for the year 2005, which was the last year for which complete information on the strains was available (serotype and subtype). The cases correspond to confirmed non-Typhi Salmonellosis cases which strains, or analysis reports, have been sent on a voluntary basis to the NRC by public and private microbiology laboratories, together with epidemiological information.

As the model is designed to attribute domestic sporadic cases, i.e. cases which are not related to an outbreak and have been infected on the national metropolitan territory, we gathered information on recent travels and outbreaks. Information regarding travelling abroad was available in the CNR database, the outbreak-related cases were identified by confronting the InVS outbreak database and the CNR database. The identified travel and outbreak related case were excluded from the dataset as well as the cases registered in the overseas departments and territories. For each outbreak, one case is kept in the dataset to represent it in the model.

2.1.2 Food-animal sources

Among the available datasets on animal food sources (David, Danan et al. submitted), we selected databases retrieving data at the national level, issued either from routine surveillance or large scale, long-term studies (at least one year). The data on the animal sources should have been collected in 2005 to match the human data, be representative and encompass at least the main channels (layers, broilers, pigs, turkeys, cattle). As the microbial subtyping approach attributes at the reservoir-level, we focused on data collected either at the farm level or at the abattoir level, as early as possible in the slaughter line.

Active surveillance

In the frame of the European regulation, and to prepare the enforcement of the zoonoses directive, one year long national prevalence studies have been conducted from 2005 to 2007 in the main animal channels either at the farm (layers, broilers, turkeys) or at the slaughterhouse level (pigs). For each food-animal species, a representative sample of flocks or carcasses has been tested for *Salmonella*, 519 layer flocks, 371 broiler flocks, 331 turkey flocks and 1 166 pig carcasses. All isolates were serotyped and the estimated prevalences were adjusted for the flock size for layers, broilers and turkeys. Data on cattle come from the national surveillance plan of antimicrobial resistance in indicator and zoonotic bacteria in cattle, a national annual monitoring plan aiming at evaluating in this channel the prevalence of resistant bacteria, among which *Salmonella*.

These data, thanks to the sampling design of the studies they are issued of, are representative of the French prevalence level at the production stage. Though, the results for broilers, turkeys and pigs are not for 2005 but for 2006 (broilers) and 2007 (turkeys and pigs). Indeed, no representative active surveillance studies were available for 2005 for those channels at the measurement points focused.

Passive surveillance

SN centralizes strains and analysis reports on non human strains of *Salmonella*, sent on a voluntary basis by a network of private and public veterinary laboratories (David, Danan et al. submitted). We extracted the 2005 serotyping data for the 5 main channels plus sheep, ducks, other poultry and sea products. The "sea products" category gathers data on fish, shellfish and seafood, the "other poultry" category gathers data on geese, guinea fowls, pheasants, quails and pigeons. We only included strains collected at the farm level, on production animals and their environment and at the abattoir level on carcasses. Strains relative to breeding animals, to diseased animals or which origin was unclear were excluded.

For cattle, an overrepresentation of pathogenic strains has been identified at the farm level. As a consequence, for this channel, we only included the strains collected at the abattoir level.

2.1.3 Subtyping

To optimize the attribution, Enteritidis and Typhimurium strains which represent more than 30% of the human cases each, have been further subtyped. The subtyping is based on antimicrobial resistance profiles (Berge, Atwill et al. 2003). For the human cases, we used the data available for 92 strains out of 3 536 for Typhimurium and on 102 out of 3 138 for Enteritidis. For the food-animal sources, the profiles realized by SN on the strains after exclusion of duplicates has been used for the passive dataset and for the active dataset, all strains of Enteritidis and Typhimurium collected in the frame of the baseline studies have been further analyzed for antimicrobial resistance especially for this study.

Antibiotic susceptibility was determined by disk diffusion on Mueller-Hinton agar according to the guidelines of the Antibiogram Committee of the French Society for Microbiology. All the routine tested antimicrobials common to the human and animal databases have been included: amoxicillin, chloramphenicol, ceftazidim, gentamicin, kanamycin, nalidixic acid, streptomycin, sulfamides, sulfamides-trimethoprim, tetracyclin. For each antimicrobial, the strains were classified either as susceptible or as not susceptible when intermediate or resistant. The subtypes were defined through Multiple Correspondence Analysis (MCA) and mixed classification (Berge, Atwill et al. 2003) applied on the obtained antimicrobial profiles.

2.1.4 Consumption

The data on food consumption used in the model come from a national survey on individual food consumption, the INCA study, conducted by the French Food Safety Agency (AFSSA) in 1999 on 3 003 representative subjects above 3 years old (Volatier 2000). We also used data

on the amount of the different meat types, eggs and sea products available for consumption, published by the French Livestock Institute and the French Interprofessional National Board for Sea and Fish-Farming Products on an annual basis. These data are converted as an annual consumption index with 1999 as reference year.

The consumption data (in tons) from the 1999 INCA study were updated with this index to obtain the 2005 consumption data. This allows having information as close as possible from the effective consumption.

2.2 Models

2.2.1 Simple model

The expected number of human cases of type i ($i=1,\dots,I$) linked to a source j ($j=1,\dots, J$) (λ_{ij}) is assumed to be proportional to the prevalence of type i in source j (p_{ij}) and to the exposition of the population to source j , measured by the amount of source j consumed by the general population (M_j). The total observed number cases of type i (o_i) is distributed among the sources in which type i is present, according to the relative weight :

$$\lambda_{ij} = \frac{p_{ij} \times M_j}{\sum_j p_{ij} \times M_j} \times o_i. \quad (1)$$

Only 3% of the human Enteritidis and Typhimurium strains have been subtyped, the 97% not subtyped have been reallocated. Similar subtype distribution of cases subtyped and not subtyped has been assumed and the observed types distribution of the cases that have been subtyped within each serotype was applied to allocate the cases with unknown subtype. This reallocated number of cases was used in place of o_i .

The expected number of cases due to source j (source $_j$) is obtained by summing the λ_{ij} on i . This simple approach assumes that all the sources are equivalent vehicles for the pathogen

and all the serotypes have the same capacity to induce infection, but allows a clearer insight into the functionality of the approach (Zwietering 2008).

2.2.2 Bayesian model

The Bayesian microbial subtyping attribution model, based on the Hald model (Hald, Vose et al. 2004), has been described in a former publication (David, informative information). As for the simple model, the number of human cases caused by different *Salmonella* types is compared with the prevalence of those types in the reservoirs, i.e. the food-animal sources, taking into account the exposition, i.e. the amount of food source consumed. Moreover, with some informative prior, the Bayesian framework allows taking the differences between sources in their capacity to vehicle the pathogen and between the types, in their capacity to cause infection, into account. The expected number of observed human cases (o_i) is assumed to be Poisson distributed:

$$o_i \sim \text{Poisson} (\sum_j \lambda_{ij}) \text{ with } \lambda_{ij} = M_j p_{ij} q_i a_j \quad (2)$$

Thus, the expected number of human cases due to a given type i in a given source j (λ_{ij}) depends on the prevalence of the type i in the source j (p_{ij}), on the consumption of the source j in the general population (M_j) and on a source-dependant and a type-dependant parameters (a_j and q_i respectively).

As defined by Hald *et al*, the type dependant factor (q_i) summarizes the characteristics of the serotype (survivability, virulence, pathogenicity, ...) which determine its capacity to cause an infection and the source dependant factor (a_j) summarizes the characteristics of the source (physical properties, preparation methods, processing procedures, ...) which determine its capacity to act as a vehicle for *Salmonella* (Hald, Vose et al. 2004). The expected number of cases due to source j ($source_j$) is obtained by summing the λ_{ij} on i .

As for the simple model, the subtypes distribution is assumed to be similar in subtyped and not subtyped cases. The cases not subtyped are reallocated within Typhimurium or

Enteritidis according to the observed subtypes proportions on a bayesian way, using Gamma distributions to reflect the uncertainty on the observed number of cases per subtype. The proportions used for the reallocation within a serotype are thus following:

$$\frac{\text{Gamma}(o_i;1)}{\sum_i \text{Gamma}(o_i;1)} \quad (3)$$

where o_i is the observed number cases of subtype i for a given serotype, Enteritidis or Typhimurium (Hald, Vose et al. 2004).

The priors were defined as we recommended (David, informative information). The type-dependant parameters relative to specific types (i.e. types only present in one food-animal source) other than Enteritidis and Typhimurium, as to avoid any interaction with the reallocation process, were fixed to the following data-based value:

$$q_i = \frac{o_i}{\sum_i o_i} \times \frac{1}{p_{ij}} \quad (4)$$

where o_i is the observed number of cases due to type i , p_{ij} is the prevalence of type i in the unique source j . We thus used the percentage of human cases divided by the prevalence in the source as an indicator for factor q . The prior distributions for the other type dependant parameters and the source dependant parameters are uniform distributions with 0 as lower value and an upper value defined as to encompass all possible values for the parameter.

This parameterization requires the availability of at least J specific types not corresponding to Enteritidis and Typhimurium, which was the case for both the active and the passive datasets. It also requires the microbiological types to be heterogeneously distributed among the sources considered.

For the passive dataset, types distributions are used in place of prevalences.

3 Results

3.1 Data

3.1.1 Human cases

The dataset included 9 076 domestic sporadic confirmed human cases. Enteritidis and Typhimurium were the most frequent serotypes, representing respectively 35% and 39% of the cases. The subtyping resulted in 9 subtypes for Enteritidis and 9 also for Typhimurium, some of which gather only human strains or only animal strains.

3.1.2 Datasets

For each dataset, serotypes or subtypes which have caused only a few (or no) infections or are not observed in the animal reservoirs are grouped into the categories "others" for serotypes and "other Typhimurium" or "other Enteritidis" for the subtypes. Subsequently, the types' composition of the "other" categories is different for the human data and for each of the sources. The relative human cases are consequently considered as not attributable and will automatically be affected to the "unknown source" category in the attribution process.

We here use two datasets with different surveillance systems for the food-sources. The number of included sources and the types observed among the food-sources strains thus differ between both datasets. As a consequence, the observed serotypes common to the human cases and the food-sources are different between the two datasets which leads for the human cases as well as for each of the common food sources to differences in the composition of the categories "Others", "Other Typhimurium" and "Other Enteritidis" in the active and in the passive dataset. For the same reasons, the number of attributable cases and the types taken into account for these attributable cases differ for both datasets.

Active surveillance dataset

5 938 human cases are attributable to the five considered food sources, spread between 28 serotypes, 5 Typhimurium subtypes and 3 Enteritidis subtypes. The global observed prevalences in the sources were 30.8% for layers, 18.8% for pigs, 13.8% for turkeys, 8.2% for broilers and 2.4% for cattle. The types distribution among the human and in the sources is presented in figure 1a. The types are heterogeneously distributed among the sources according to Fisher exact tests.

Among the ten most frequent types within the attributable human cases, 6 are Typhimurium or Enteritidis subtypes. Overall, Enteritidis and Typhimurium represent over 75% (respectively 47.1 and 30.4%) of the human attributable cases. The human cases are mostly associated to SE-multiS (35.2%), ST1 (17.5%), SE1 (10.4%), ST4 (6.5%) and Infantis (2.8%).

The food-source cattle is characterized by a small number of types, 2 when excluding the "others" category, namely Indiana and Mbandaka, this last serotype representing 62.5% of all cattle strains. Layers are mainly associated to Infantis (13.2% of the strains), Anatum (12.6%), Bareilly (7.1%), ST5 (6.8%) and Mbandaka (6.8%). For pigs, turkeys and broilers, three serotypes each appear to be much more frequent than the others: Derby (35.9%), ST1 (19.8%) and ST5 (13.2%) for pigs ; Derby (25.3%), ST3 (13.4%) and Mbandaka (11.2%) for Turkeys ; Hadar (17.7%), Virchow (17.4%) and Montevideo (14.9%) for broilers. Thus, Derby appears to be the most frequent serotype both in the turkeys and broilers sources. Moreover, when pooling Enteritidis and Typhimurium subtypes (excluding the "other" categories), Enteritidis is mainly observed in the layers source, where it represents 7.9% of the strains. Typhimurium, if ubiquitous, is mostly associated to pigs (38.7% of the strains) and to a less extent to layers (14.3%) and to broilers (13.4%).

This dataset globally comprises 14 specific types (table 1), including one Typhimurium subtype and one Enteritidis subtype. All sources but cattle are represented among the specific types, pigs and layers being the most frequently concerned.

Passive surveillance dataset

6 527 human cases are attributable to the 9 included sources, spread between 37 serotypes, 5 Typhimurium subtypes and 4 Enteritidis subtypes. The proportion of the types within a food-animal species are given for 617 strains for layers, 1 273 strains for broilers, 796 strains for turkeys, 213 strains for pigs, 71 strains for cattle, 104 strains for sheep, 40 strains for the sea products, 3 113 strains for ducks and 722 strains for other poultry. The types distribution among the human cases and in the sources is presented in figure 1. The types are heterogeneously distributed among the sources according to Fisher exact tests.

Among the ten most frequent types within the attributable human cases, 7 are Typhimurium or Enteritidis subtypes. Overall, Enteritidis and Typhimurium represent over 75% (respectively 47.6 and 27.7%) of the human attributable cases. The five main types are almost the same as for the active dataset, that is SE-multiS (32.1%), ST1 (15.9%), SE1 (9.4%), ST4 (5.9%) and SE3 (4.7%).

Layers, broilers, other poultry and turkeys are characterized by the diversity of serotypes and subtypes observed (29, 27, 27, 26 respectively). Layers and broilers are especially associated to Senftenberg (19.8% and 13.0% of the strains respectively) and SE-multiS (16.9% and 16.6% respectively). For turkeys, the most frequent serotype is Senftenberg (14.7%). Two serotypes emerge for other poultry: Indiana (31.0%) and ST2 (14.5%). Serotype Indiana is also the most frequent serotype in ducks (23.1%). In the passive dataset, the types associated to cattle are more diverse, 11 serotypes (vs 2 in the active dataset), the most frequent ones being ST1 (22.5%), Derby (19.7%) and Montevideo (14.1%). Sea products are mostly associated to Weltevreden (25%). At last, sheep is the source with the smallest number of types, that is 7. This source is particularly associated to the serotype S 61:k:1,5,7, which gathers 57.7% of the strains.

When pooling Enteritidis and Typhimurium subtypes (excluding the "other" categories), Enteritidis is mostly found in layers (24.3%) and broilers (16.7%), whereas Typhimurium is

rather associated to other poultry (30.7%), pigs (23.5%), cattle (22.5%) and sea products (15.0%).

This dataset comprises 15 specific types (table 1), including two Enteritidis subtypes. As for the active surveillance, all sources except cattle are represented among the specific types and pigs and layers are the most frequently concerned, in number of specific types and in number of associated human cases. Four specific types are common for both samples: SE2, Oranienburg and Havana, linked to layers and Goldcoast linked to pigs.

Thus, the included serotypes, even if globally similar, are not exactly the same in both datasets. This leads to a difference in the number of attributable cases, which is 10% higher in the passive dataset. When considering the 5 common food-animal species, the types distribution appears different between the two datasets for a given species, which was confirmed by Fisher exact tests. However, in both samples, Derby is the most frequent serotype for the pigs, which is also true for Typhimurium globally. And SE-multiS is the most frequent type in the layers whichever the source of the data is. Specific types were identified for all food sources except cattle in both datasets, though there are few common ones which concern only layers and pigs. Actually, it appears that each type of surveillance gives a different "image" of the food-sources contamination, with significantly different types repartitions for common food sources.

3.1.3 Consumption data

On the basis of the updated results of the INCA study, we assumed that the national consumption for 2005 was 82 301 tons for layers (eggs), 84 842 tons for broilers, 18 967 tons for turkeys, 161 971 tons for pigs, 144 504 tons for cattle, 20 232 tons for sheep, 111 764 tons for sea products, 5 309 tons for ducks and 4 580 tons for other poultry. Here appears that the sea products is a major potential source in term of consumption, though little information is available on the *Salmonella* contamination of such products.

3.2 Impact of the data characteristics on the attribution results using the simple model

To analyze the impact of the data characteristics on the attribution results, we used the simple model in order to benefit from a clear insight in its functionality.

3.2.1 Contamination of the sources: proportion vs prevalence

For the passive dataset, prevalences are not available, but only strains proportions per type. We here assessed the impact of using proportions instead of prevalences with the active surveillance dataset, for which both indicators were available (figure 2).

No change in the ranking of the sources was observed, but the numbers of cases attributed to each source differed. For the most important source, layers, using proportions led to a decrease of 25.7% (3 164 vs 2 350 cases respectively). Inversely, an increase in the number of attributed cases was observed for the four other food-sources, 8.2% (1 901 vs 2 057) for pigs, which is the second most important source and 84.4% for broilers (925 vs 501), 73.5% for cattle (62 vs 36) and 62.2% for turkeys (544 vs 336). For cattle, though, the number of attributed cases remained modest.

Thus, it appears possible to use proportions instead of prevalences, though, in this case the numbers of cases per source have to be interpreted with caution and the relative importance of the sources seems to be more reliable.

3.2.2 Increasing sources' number

The passive dataset allows studying the effect of including more sources. The model was first estimated with the five main sources, common with the active dataset, then with the nine available sources.

When including nine sources, the total number of attributable cases increased slightly, by 1.2% (figure 3). For each of the main sources, the number of attributed cases decreased, from 2.1% for layers to 12.5% for pigs, but the relative importance of those 5 sources remained the same. The cases attributed to the 4 additional sources thus correspond partly to newly attributed cases (1.2%) and partly to reattributed cases (around 7%) which were probably wrongly attributed when considering only five sources.

Thus, including additional sources allows attributing more cases and avoiding some misattributions, without changing the relative importance of the initially included sources.

Finally, comparing the number of attributed cases for the main sources between active and passive samples, using the prevalences for the active dataset (figures 2 and 3), two major differences appear which concern broilers and cattle. First, with the passive dataset, broilers is the second most important source, after layers, whereas with the active dataset broilers was at the third place with a much lower number of attributed cases. Namely, an increase of 171% is observed with the passive data (1 359 cases vs 501 cases). Second, cattle which is the less important source by far when considering the active dataset (36 cases) becomes the third source when using the passive dataset, with a quite high number of associated cases (1 170 cases).

When examining the impact of the data characteristics on the attribution results obtained with a simple model, it appeared that using prevalences or proportions had no major impact on the attribution results, as did the inclusion of additional sources. The important increase in the number of cases associated to broilers and cattle when using passive surveillance data is probably linked to the observed serotypes. Both sources present namely differences in the types repartition for the common sources. And as the representativeness is not assured for the passive data on food source contamination, interpreting results based on active surveillance data seems more convenient.

3.3 Attribution results using Bayesian model

The bayesian model is applied separately with the active and passive datasets. Results presented correspond to runs of 100 000 iterations of the Gibbs sampler with a thin of 25. Convergence diagnostics were satisfactory (Cowles and Carlin 1996; Brooks and Gelman 1998; Brooks and Roberts 1998b). From these runs, parameters estimates (posterior means, posterior variances and posterior 95% credibility intervals (95% CI) were computed from the last 50 000 iterations.

The expected numbers of cases per type and the observed numbers of cases, or the deterministically reallocated numbers of cases for Enteritidis and Typhimurium subtypes, are compared. They show a good adequacy for both datasets (Figure 4).

The total expected number of cases is, as expected, higher with the passive surveillance, 7 122 (95% CI: 6 719–7 528) compared to 5 746 (95% CI: 5 307 – 6 172) with the active surveillance dataset. For the active surveillance, the expected number of cases is in agreement with the 5 938 attributable cases, however, for the passive surveillance, it is slightly overestimated in regard to the 6 527 attributable cases.

For the 5 main sources, the numbers of expected cases per source are not significantly different when using the passive dataset, except for broilers (figure 5). It is lower for layers (2 328 vs 3 060), turkeys (336 vs 714) and pigs (1 039 vs 1 477) and higher for cattle (133 vs 69). For broilers, the number of cases is significantly higher with the passive dataset (1 410 vs 427) which was also observed with the simple model. This is probably linked to the higher frequency in this source of Typhimurium and Enteritidis subtypes in the passive dataset (27.4% globally) than in the active dataset (5.1% of the strains). Moreover, unlike when using the simple model, the number of cases attributed to cattle isn't significantly different for both datasets. This confirms the interest of the bayesian modelling and the relevance of introducing source and type dependant parameters. Actually, the source

dependant parameter relative to cattle is significantly smaller for the passive dataset (0.0039, 95% CI: 6.6E-5 – 0.018 vs 13.5, 95% CI: 0.025 – 83.6 for the active dataset), whereas, for the other common food-sources the posterior means are not significantly different (Table 2).

Except for cattle, the source dependant factors are not significantly different between the two datasets for the main sources. Even if not significantly, the best vehicle among those four sources for *Salmonella* is turkey meat, which also has the widest 95% CI. The other sources are very close. No firm conclusion can be drawn for cattle, which source dependant factor is the highest for the active dataset and the smallest for the passive dataset, with very wide 95% CI. Finally, in the passive dataset, the highest source dependant factors are relative to ducks (1.37) and other poultry (1.06), the difference with the other sources being significant. The better vehicle for *Salmonella* thus appears to be birds, which are also especially linked to Enteritidis, the most frequent serotype among the attributable human cases.

Concerning the type-dependant parameters, the posterior means obtained with the passive dataset are systematically lower than the ones obtained with the active dataset, except for Stourbridge, mostly significantly. Another global feature is that the 95% CI are less wide for Enteritidis and Typhimurium subtypes with the passive dataset. When considering the 10 highest type dependant parameters for each dataset (table 3), 6 types are common: SE1, SE-multiS, Stourbridge, Havana, Oranienburg and Napoli, but only the parameter relative to Stourbridge has the same value in both dataset, which is a fixed value in the active dataset (5.40) and an estimated value in the passive dataset (5.41 95% CI: 2.42 – 10.72). For Napoli, if the values are different, the rank is the same in both cases: 10. It can also be observed that 3 highly ranked parameters in the active dataset (corresponding to Heidelberg, ST2 and SE2) are low ranked in the passive dataset, respectively 26th, 33rd and

19th rank, instead of 1st, 3rd and 8th rank. For what concerns the passive dataset, 4 types in this top ten are not observed in the active dataset, all the parameters relative to the other types, are all also in the top ten in the active dataset. At last, when considering Enteritidis and Typhimurium relative parameters, it can be seen that all Enteritidis subtypes but one, are in the top ten for each dataset, which underlines the infectious capacity of this serotype, and the only Typhimurium subtype appearing here is ST2 in the active dataset, which gathers multi-resistant strains resistant to nalidixic acid.

To summarize the results, with the active surveillance dataset, the major source of infection is layers (eggs) to which 53.3% of the expected cases are attributed (95% CI: 46.0-60.0), followed by pigs, to which 25.7% (95% CI: 20.7-31.1) of the cases are attributed. The third place is taken by broilers and turkeys which results are not significantly different, respectively 7.4% and 12.4% (95% CI: 5.6-9.5 and 6.9-19.2). Cattle appears to be a minor source, which comprised only 1.2% of the cases (95% CI: 0.2-1.7). The cases which origin remains unknown are as numerous as the expected number of cases attributed to layers.

Using the passive dataset leads to comparable results concerning layers, pigs, turkeys and cattle, though for broilers, the number of attributed cases is significantly higher (1 410, 95% CI 1 026-1 821), so that this source is comparable to pigs (1 039, 95% CI: 785-1 310) in importance. The ranking of the sources so would be layers as the most important source of infection, broilers and pigs as second importance sources, all the other sources (turkeys, cattle and the four additional) following, without distinction as the expected numbers of attributed cases are near from each other and the 95% CI quite wide. As for the results obtained with the active dataset, the cases of unknown origin are as numerous as the number of cases attributed to layers.

Concerning a and q parameters, it is difficult to draw any firm conclusion on the comparative infectious capacities of the different serotypes as well as on the relative capacity of the sources to vehicle the pathogen. Namely, when some parameters are higher than the others,

their 95% CI are wide. Moreover, the results even in term of ranking diverge between both datasets. Though, it can be extrapolated from the few convergent results, that birds appear to be a good vehicle for *Salmonella* and that Enteritidis has a high capacity of infection, whichever the subtype is, as do some serotypes such as Stourbridge, Havana, Oranienburg and Napoli.

4 Discussion

The application of the microbial subtyping approach as conceived by Hald *et al* requires intensive monitoring data on human cases and all relevant food-sources, moreover the contamination of sources should be measured in prevalence and discriminatory typing methods, beyond serotyping, should be applied (Hald, Vose et al. 2004). Producing representative prevalence data at the farm and at the abattoir levels is essential. It necessitates to actively monitor the main food-sources, which is costly and not always developed for this early point of the food-chain (Batz, Doyle et al. 2005; Kirk, McKay et al. 2008; Sofos 2008). Thus, passively acquired surveillance data can be the only available alternative, even if their representativeness is not assured. In France, thanks to the European regulatory frame and to long lasting lab-based surveillance systems, all the required data were available and the typing methods were harmonized for all datasets, as a consequence of the ongoing cooperations for outbreaks investigations. Moreover, we disposed for the sources both of representative prevalence data and of passive surveillance data, which allowed us assessing the relative advantages of using such data in the bayesian attribution approach in terms of results quality and of potential sources coverage.

A large scope collaboration, which is a key point for such projects (Batz, Doyle et al. 2005; ICMSF 2006), allowed centralizing all the required data, which were fragmented between several actors. This led to the build an active and a passive dataset. The heterogeneity in the repartition of the types among the sources was met, as did the necessary number of specific types for the parameterization of the model. Though, we had to deal with some insufficiencies in the data retrieved. The human database is produced by a lab-based surveillance network, thus neither exhaustive, nor definitely representative. Moreover, the information on travel was not always accurate. Though, as it relies on a stable and nationwide network of public and private laboratories, which representativeness is estimated as good, it constitutes a reasonable basis for an attribution work. Another problem is the lack of time consistency of the actively acquired food-source data with the human cases. This raises concern about a potential bias on the results obtained with the active dataset. Though, this will be solved for the coming years, as the main sources (layers, broilers, turkeys and pigs) will be regularly monitored as required by the European Zoonosis Directive which is actually enforced (David, Danan et al. submitted). Finally, the subtyping of human Enteritidis and Typhimurium strains concerned only 3% of the cases, the other cases being reallocated in the model, which led to wide credibility intervals for the type-dependant parameters and contributed to the uncertainty on the attribution results. Moreover, assuming similar distribution of cases subtyped and not subtyped seems in these conditions questionable. Testing more human strains for antimicrobial resistance would certainly enhance the predictions of the model. We thus disposed of a complete dataset, based on harmonized typing methodologies, but, because of some imperfections, and especially of the absence of any data on the imported products and a lack of time consistency of the active data, the attribution results presented, if valuable to study the methodological points (passive vs active data), have to be taken with caution regarding their representativeness and should be enhanced in future studies.

The main point studied was the comparison of the results obtained with passive and active data regarding the source contamination in the model. First of all, the image in term of types distribution given by the both surveillance systems for the common sources were significantly different. Moreover, the data characteristics were different in term of contamination indicator and of number of included sources. Using the simple model, these characteristics had little impact on the attribution results when considered singly, but when comparing the results obtained for the differences altogether, they were quite different (especially for broilers and cattle). The bayesian model, on the contrary, seems to be robust to this heterogeneity, as the results are in accordance for both datasets, except for broilers. Thus, passive data appear to be a potential complement for active datasets. Namely, the active and systematic surveillance of food sources is costly and thus can only encompass a limited number of sources. Including additional sources through the passive surveillance has several advantages. As one of the assumptions of the model is that all the attributable cases are linked to the included sources, including as many potential sources as possible allows avoiding misattributions. Moreover, as the contamination of the sources at the farm level evolves according to public health interventions, enforcement of regulations and changes in hygiene practices (Rostagno, Hurd et al. 2005; Denagamage, O'Connor et al. 2007), and as the exposition of the population to the sources also changes (Combris 1997; Allard 2002; Desenclos, Vaillant et al. 2002) the model can help detecting the emergence as public health concerns, of sources not institutionally monitored but which should be. Finally, it is cost-effective, which is an important point in public health tools design (Havelaar, Bräunig et al. 2007). Thus, when applied with both active and passive data, the model could be a useful tool to contribute to the adjustment of the institutional active surveillance, besides its potential role for the conception and the evaluation of public health interventions.

If only passive surveillance data are available for the sources contamination, reasonably valuable results are obtained. Of course in this case, conclusions will have to be made with caution, taking the quality of the data into account.

The results obtained on the relative importance of the sources, when considering the active dataset complemented by the passive dataset for the 4 additional sources, are in accordance with the Danish results (Hald, Vose et al. 2004; Havelaar, Bräunig et al. 2007): layers being the most important source, followed by pigs and broilers, turkeys, ducks and cattle which are not significantly different. Broilers is the only source for which the expected number of cases is significantly different between both datasets. This could be due to the higher frequency of Typhimurium and Enteritidis among the strains in the passive dataset, the source-dependant factor being similar between both datasets. Such a difference in the results obtained for *Salmonella* with active and passive data especially for broilers (compared to pigs) has been recently exposed in a different area (antimicrobial resistance) (Mather, Haydon et al. 2009). Thus, it would be interesting to explore the reasons of this difference, specific to broilers.

For each configuration, the number of cases with unknown origin was as high as the number of expected cases attributed to the main source, i.e. layers. The major part of these "unknown" cases belonged to Enteritidis and Typhimurium "other" categories, corresponding to subtypes not observed in the considered animal sources. An explanation for this could be either a bias due to the small proportion of strains subtyped among the human cases or a detection problem of rare types in the sources. Though, in this last case, the proportion of unknown cases should be much lower with the passive dataset, as the passive surveillance allowed gathering more strains for each common animal species and to include more sources and thus potentially better detect rare types. Another probable reason for this high frequency of cases with unknown reservoir, is to be found in the potential sources not considered in this analysis. We namely only considered the foodborne transmission, thereby excluding sources such as pets, and exotic pets which are known sources of human Salmonellosis (Woodward, Khakhria et al. 1997; Bellido Blasco, Gonzalez Cano et al. 1998; Mermin, Hutwagner et al. 2004; De Jong, Andersson et al. 2005; Finley, Reid-Smith et al. 2006; Marcus, Varma et al. 2007) or even man to man transmission (Todd, Greig et al. 2008). Moreover among the potential food sources, we focused on some animal food-

sources, which are certainly the most frequent sources of human Salmonellosis but not the only ones. Plants (Brandl 2006; Elviss, Little et al. 2009) and other non animal products (Kirk, McKay et al. 2008) have been implicated as a source for Salmonellosis. At last, imported products, though representing from 15% (for eggs) to 60% (Lamb) of the national consumption according to the food-animals products, could not be taken into account in this analysis but also are potential sources for these unknown cases.

Direct comparison between the source dependant factors for the two datasets was difficult. Indeed, the contamination indicator is prevalence in one case and proportion in the other case and as a consequence, the source dependant parameter doesn't measure exactly the same thing, which is also true for the type dependant factor. Actually, the type dependant factors were almost systematically different when using one or the other dataset, even when the ranking was coherent. The estimated source dependant factors though, were in good agreement between both datasets, except for cattle. Moreover, the 95% CI were mostly too wide to obtain significant differences.

Hald et al underlined that these factors were only multiplication factors allowing arriving at the most probable solution (Hald, Vose et al. 2004). This was actually the case in our study, especially when considering cattle, for which the attribution results diverged using the simple model, but were in accordance when introducing the a and q parameters. Though, it was difficult to draw any firm conclusion on the obtained posterior distributions, especially for the type dependant parameter. Actually, those parameters are conceived as black boxes, they allow taking into account differences between types and between sources, which is of primary importance (Blaser and Newman 1982; D'Aoust 1989; Sarwari, Magder et al. 2001; Coleman, Marks et al. 2004; Bollaerts, Aerts et al. 2008; Jones, Ingram et al. 2008), but their nature is not defined. It thus would be interesting to give a definition to these parameters, which would allow to use exogenous information, such as dose-response relationship, infective dose, pathogenicity (Bollaerts, Aerts et al. 2008; Jones, Ingram et al. 2008),..., to validate informative prior distributions.

Conclusion

We here addressed the problem of the data quality, taking advantage of the existence in France of an active and a passive surveillance system for what concerns the contamination of the sources. Our results indicate that it is possible to obtain reasonably reliable attribution estimates when using passively acquired data and even that those data can be an interesting and cost-effective complement to the active and representative prevalence data. Though, the results for the additional sources will remain indicative and conclusions should be drawn on the ranking rather than on the absolute number of cases. Moreover, the sources and type dependant parameters can hardly be interpreted, given the differences between both datasets results. For France, the improvement of the approach should be obtained by an improvement of the human and animal datasets, for example by a better collection of data about travel history of the human cases, the sub-typing a larger subset of Enteritidis and Typhimurium human strains and the collection of data about prevalence of salmonella in main imported animal sources. These recommendations applied to further studies would lead to a decrease of uncertainty around estimates. The Bayesian tool could also be improved by a research work about the nature and dimension of the type and source parameters which are of primarily importance in the approach. Finally, the parallel use of active and passive data could contribute to a risk-based orientation of the active surveillance policy.

References

- Allard, D. G. (2002). "The 'farm to plate' approach to food safety - Everyone's business." Canadian Journal of Infectious Diseases **13**(3): 185-190.
- Batz, M. B., M. P. Doyle, et al. (2005). "Attributing illness to food." Emerging Infectious Diseases **11**(7): 993-999.
- Bellido Blasco, J. B., J. M. Gonzalez Cano, et al. (1998). "Factors associated with sporadic cases of salmonellosis in 1- to 7-year-old children. Study of cases and controls." Factores asociados con casos esporádicos de salmonellosis en niños de 1 a 7 años. Estudio de casos y controles. Gaceta sanitaria / S.E.S.P.A.S **12**(3): 118-125.
- Berge, A. C. B., E. R. Atwill, et al. (2003). "Assessing antibiotic resistance in fecal *Escherichia coli* in young calves using cluster analysis techniques." Preventive Veterinary Medicine **61**(2): 91-102.
- Blaser, M. J. and L. S. Newman (1982). "A review of human salmonellosis: I. Infective dose." Reviews of Infectious Diseases **4**(6): 1096-1106.
- Bollaerts, K., M. Aerts, et al. (2008). "Human salmonellosis: estimation of dose-illness from outbreak data." Risk Anal **28**(2): 427-40.
- Brandl, M. T. (2006). "Fitness of human enteric pathogens on plants and implications for food safety." Annu Rev Phytopathol **44**: 367-92.
- Brooks, S. P. and A. Gelman (1998). "General methods for monitoring convergence of iterative simulations." Journal of Computational and Graphical Statistics **7**(4): 434-455.
- Brooks, S. P. and G. O. Roberts (1998b). "Convergence assessment techniques for Markov chain Monte Carlo." Statistics and Computing **8**(4): 319-335.
- Coleman, M. E., H. M. Marks, et al. (2004). "Discerning strain effects in microbial dose-response data." J Toxicol Environ Health A **67**(8-10): 667-85.
- Combris, P. (1997). "The consumption of animal products in France : current tendencies and future perspectives." Productions Animales **10**(4): 267-274.
- Cowles, M. K. and B. P. Carlin (1996). "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review." Journal of the American Statistical Association **91**(434): 883-904.
- D'Aoust, J. (1989). *Salmonella. Foodborne Bacterial Pathogens*. M. P. Doyle. New York, Marcel Dekker: 327-445.
- David, J. M., C. Danan, et al. (submitted). "Structure of the French farm to table surveillance system for *Salmonella*."
- De Jong, B., Y. Andersson, et al. (2005). "Effect of regulation and education on reptile-associated salmonellosis." Emerging Infectious Diseases **11**(3): 398-403.
- Denagamage, T. N., A. M. O'Connor, et al. (2007). "Efficacy of vaccination to reduce *Salmonella* prevalence in live and slaughtered swine: A systematic review of literature from 1979 to 2007." Foodborne Pathogens and Disease **4**(4): 539-549.
- Desenclos, J. C., V. Vaillant, et al. (2002). "[Food-borne infections: do we need further data collection for public health action and research]." Rev Epidemiol Sante Publique **50**(1): 67-79.
- Edel, W. (1994). "Salmonella enteritidis eradication programme in poultry breeder flocks in The Netherlands." International Journal of Food Microbiology **21**(1-2): 171-178.
- Elviss, N. C., C. L. Little, et al. (2009). "Microbiological study of fresh herbs from retail premises uncovers an international outbreak of salmonellosis." Int J Food Microbiol.
- Finley, R., R. Reid-Smith, et al. (2006). "Human health implications of *Salmonella*-contaminated natural pet treats and raw pet food." Clin Infect Dis **42**(5): 686-91.

- Hald, T., D. M. A. Lo Fo Wong, et al. (2007). "The attribution of human infections with antimicrobial resistant Salmonella bacteria in Denmark to sources of animal origin." Foodborne Pathogens and Disease **4**(3): 313-326.
- Hald, T., D. Vose, et al. (2004). "A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis." Risk Anal **24**(1): 255-69.
- Havelaar, A. H., J. Bräunig, et al. (2007). "Towards an integrated approach in supporting microbiological food safety decisions." Zoonoses and Public Health **54**(3-4): 103-117.
- ICMSF, I. C. o. M. S. f. F. (2006). "Use of epidemiologic data to measure the impact of food safety control programs." Food Control **17**(10): 825-837.
- InVS (2004). Morbidité et mortalité dues aux maladies infectieuses d'origine alimentaire en France: 192.
- Jones, T. F., L. A. Ingram, et al. (2008). "Salmonellosis outcomes differ substantially by serotype." J Infect Dis **198**(1): 109-14.
- Kirk, M. D., I. McKay, et al. (2008). "Food safety: foodborne disease in Australia: the OzFoodNet experience." Clin Infect Dis **47**(3): 392-400.
- Marcus, R., J. K. Varma, et al. (2007). "Re-assessment of risk factors for sporadic Salmonella serotype Enteritidis infections: A case-control study in five FoodNet Sites, 2002-2003." Epidemiology and Infection **135**(1): 84-92.
- Mather, A. E., D. T. Haydon, et al. (2009). Comparing active and passive surveillance of antimicrobial resistance: Are we getting the whole story? 12th International Symposium on Veterinary Epidemiology and Economics, Durban.
- Mermin, J., L. Hutwagner, et al. (2004). "Reptiles, amphibians, and human Salmonella infection: A population-based, case-control study." Clinical Infectious Diseases **38**(SUPPL. 3).
- Mullner, P., G. Jones, et al. (2009). "Source attribution of food-borne zoonoses in New Zealand: A modified hald model." Risk Analysis **29**(7): 970-984.
- Pires, S. M., E. G. Evers, et al. (2009). "Attributing the human disease burden of foodborne infections to specific sources." Foodborne Pathog Dis **6**(4): 417-24.
- Poirier, E., L. Watier, et al. (2008). "Evaluation of the impact on human salmonellosis of control measures targeted to Salmonella Enteritidis and Typhimurium in poultry breeding using time-series analysis and intervention models in France." Epidemiol Infect **136**(9): 1217-24.
- Rostagno, M. H., H. S. Hurd, et al. (2005). "Resting pigs on transport trailers as an intervention strategy to reduce Salmonella enterica prevalence at slaughter." Journal of Food Protection **68**(8): 1720-1723.
- Sarwari, A. R., L. S. Magder, et al. (2001). "Serotype distribution of Salmonella isolates from food animals after slaughter differs from that of isolates found in humans." Journal of Infectious Diseases **183**(8): 1295-1299.
- Sofos, J. N. (2008). "Challenges to meat safety in the 21st century." Meat Science **78**(1-2): 3-13.
- Todd, E. C., J. D. Greig, et al. (2008). "Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 5. Sources of contamination and pathogen excretion from infected persons." J Food Prot **71**(12): 2582-95.
- Volatier, J. L. (2000). Enquête nationale sur les consommations alimentaires. Tec&Doc, CREDOC-AFSSA-DGAL.
- Wegener, H. C., T. Hald, et al. (2003). "Salmonella control programs in Denmark." Emerging Infectious Diseases **9**(7): 774-780.
- Woodward, D. L., R. Khakhria, et al. (1997). "Human salmonellosis associated with exotic pets." Journal of Clinical Microbiology **35**(11): 2786-2790.
- Zwietering, M. H. (2008). "Quantitative risk assessment: Is more complex always better? Simple is not stupid and complex is not always more correct." Int J Food Microbiol.

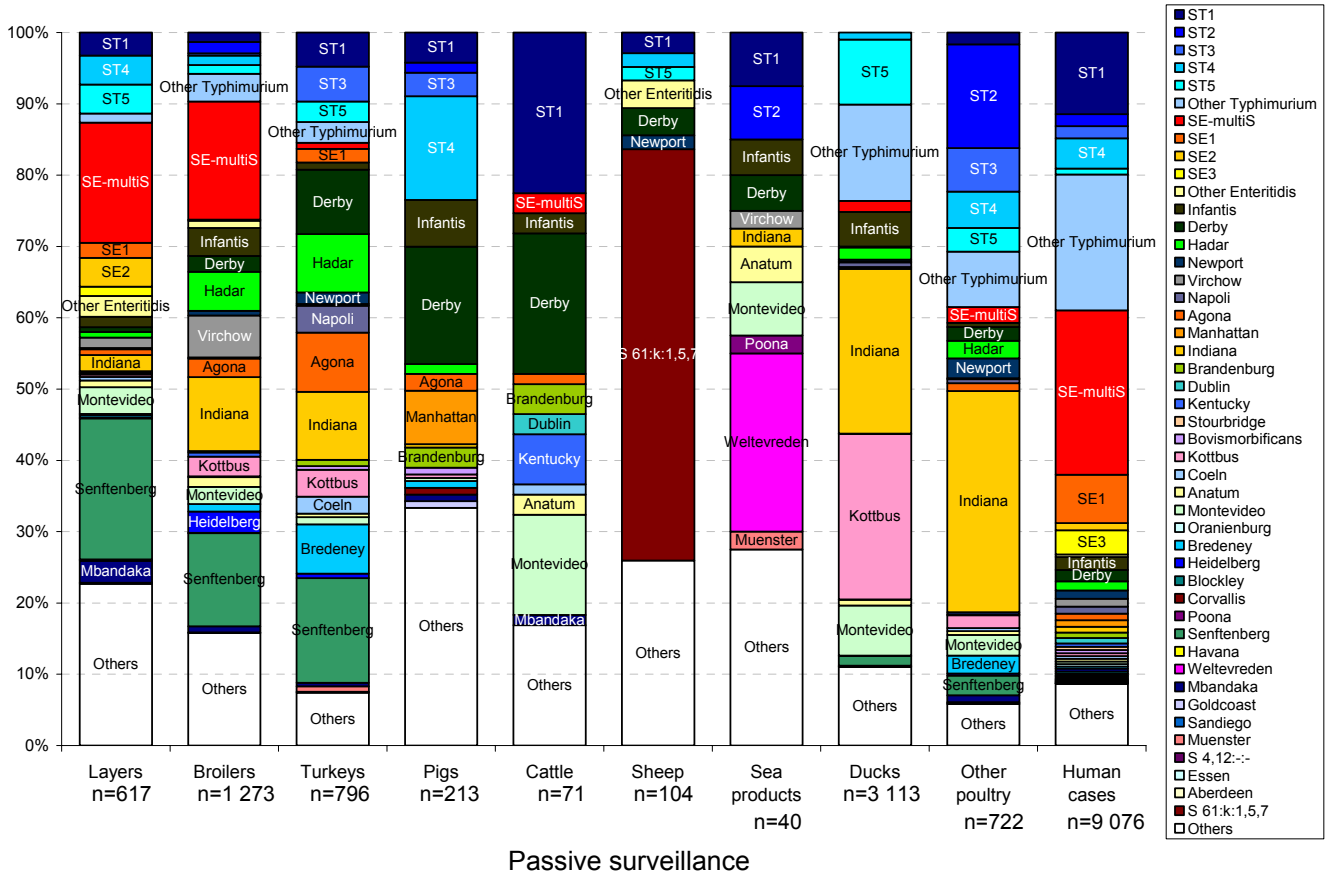
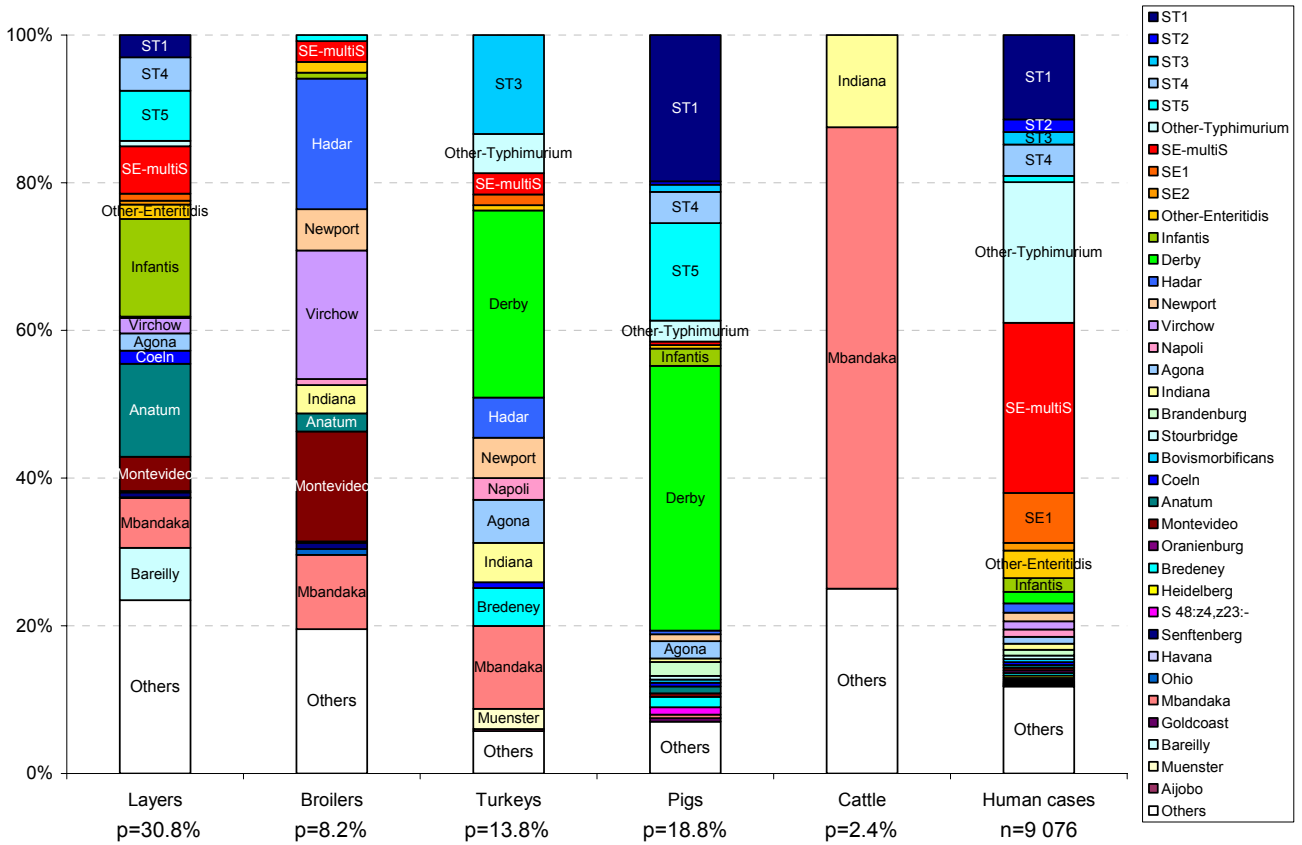


Figure 1: type distribution in the human cases and the sources according to active and passive surveillance
 p: overall prevalence in the source

Active surveillance				Passive surveillance			
Type	Source	Prevalence (%)	Number of cases	Type	Source	Proportion (%)	Number of cases
Heidelberg	Broilers	0.02	30	Blockley	Broilers	0.08	29
Ohio	Broilers	0.07	12	SE3	Layers	1.30	308
SE2	Layers	0.20	92*	SE2	Layers	4.05	92*
Oranienburg	Layers	0.09	31	Oranienburg	Layers	0.16	31
Havana	Layers	0.04	16	Havana	Layers	0.16	16
Bareilly	Layers	2.95	8	Manhattan	Pigs	7.51	81
ST2	Pigs	0.09	154*	Corvallis	Pigs	0.94	24
Goldcoast	Pigs	0.09	9	Goldcoast	Pigs	0.94	9
Stourbridge	Pigs	0.09	42	S 4,12:-:-	Other poultry	0.14	4
Bovismorbificans	Pigs	0.09	40	Aberdeen	Other poultry	0.14	2
S 48:z4,z23:-	Pigs	0.18	24	Poona	Sea products	2.50	21
Brandenburg	Pigs	0.34	71	Weltevreden	Sea products	25.0	9
Muenster	Turkeys	0.38	7	Essen	Turkeys	0.13	2
Aijobo	Turkeys	0.04	2	Sandiego	Ducks	0.06	8
				S 61:k:1,5,7	Sheep	57.7	1

Table 1: Characteristics of the specific types
* : reallocated number of cases

	Active dataset			Passive dataset		
	Mean	95CI		mean	95CI	
Layers	0.094	0.070	0.12	0.097	0.070	0.13
Broilers	0.092	0.066	0.12	0.090	0.059	0.13
Turkeys	0.32	0.14	0.59	0.11	0.02	0.31
Pigs	0.054	0.046	0.062	0.055	0.045	0.065
Cattle	13.5	0.025	83.64	0.0039	6.57E-05	0.018
Sheep	-	-	-	0.54	0.056	1.62
Sea products	-	-	-	0.076	0.050	0.11
Ducks	-	-	-	1.37	0.59	2.51
Other poultry	-	-	-	1.06	0.38	2.16

Table 2 : Posterior means and 95% CI of the source-dependant parameters

	Active dataset				Passive dataset			
	mean	95CI		Ranking	mean	95CI		Ranking
Heidelberg	19.24	-	-	1				26
ST2	14.01	0.65	36.75	3				33
Bovismorbificans	5.14	-	-	6				13
SE2	4.26	0.17	12.82	8				19
SE1	14.16	8.45	21.67	2	2.88	1.61	4.56	5
SE-multiS	8.18	6.17	10.56	4	0.73	0.55	0.94	9
Stourbridge	5.40	-	-	5	5.41	2.42	10.72	1
Havana	4.41	-	-	7	1.09	-	-	8
Oranienburg	3.89	-	-	9	2.11	-	-	6
Napoli	3.23	1.65	5.80	10	0.56	0.26	0.98	10
Blockley				NO	4.07	-	-	2
Dublin				NO	3.19	0.77	6.26	3
SE3				NO	3.03	1.45	5.41	4
Sandiego				NO	1.37	-	-	7

Table 3: Top 10 of the estimated type dependant parameters for each dataset
NO: not observed in the dataset

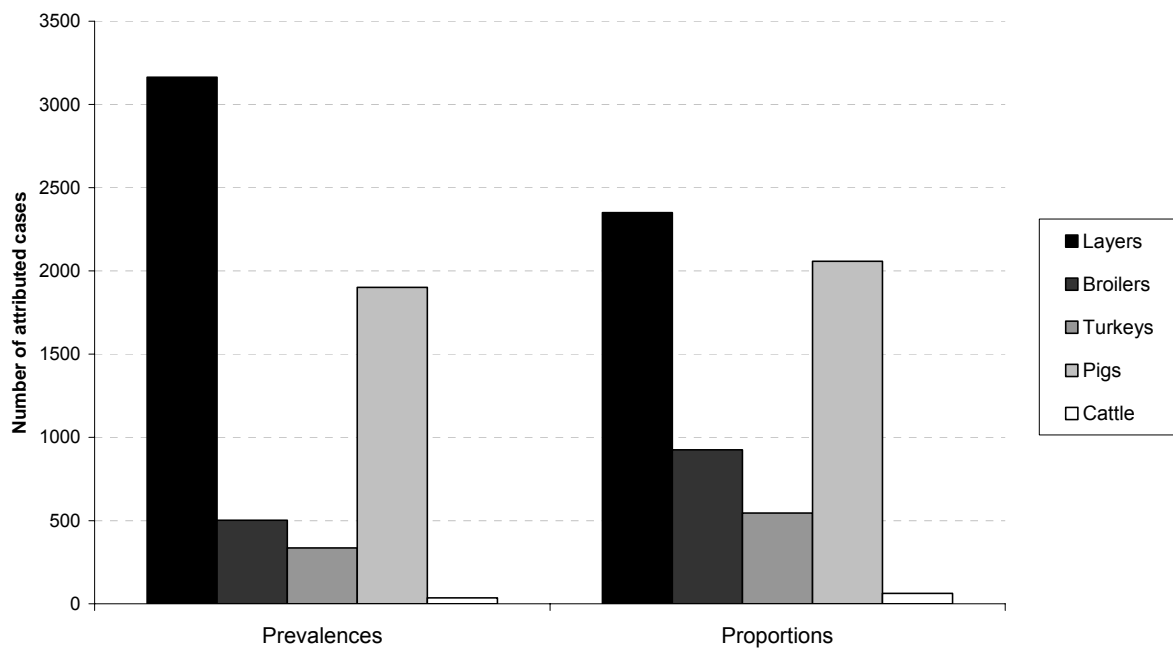


Figure 2: Impact of the indicator on the attribution results, simple model, active dataset

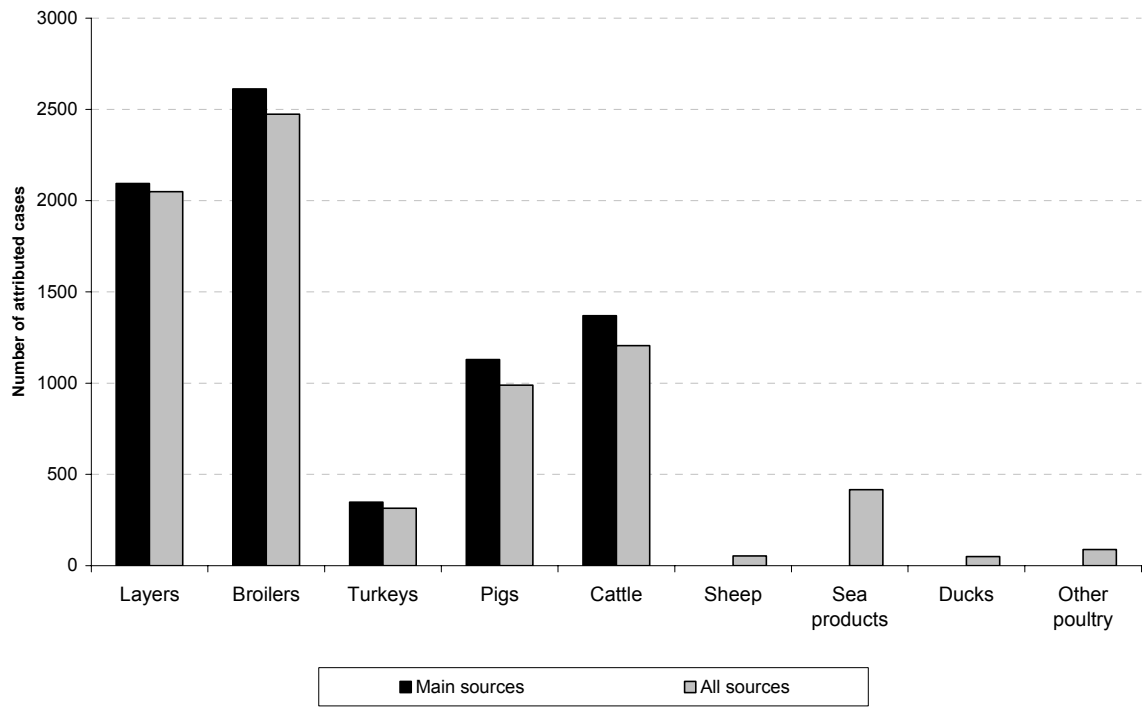


Figure 3: Impact of the number of sources included in the model on the attribution results, simple model, passive dataset

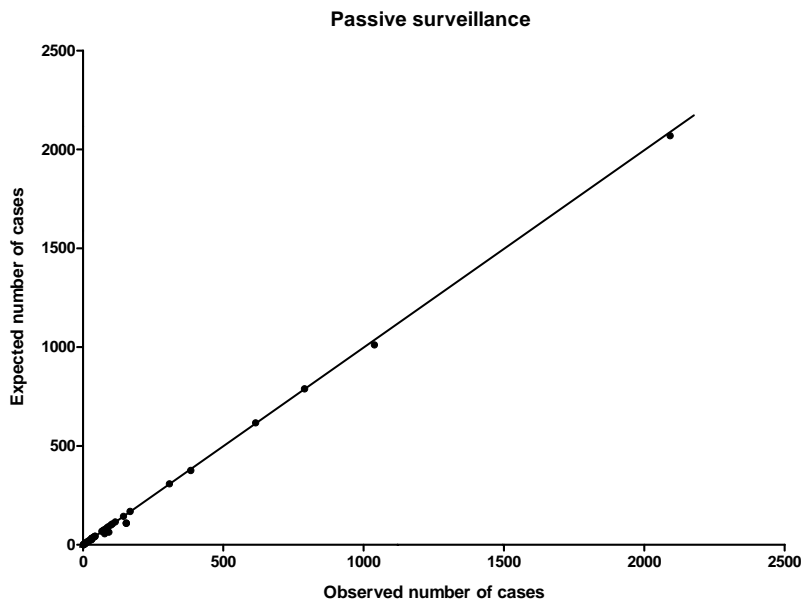
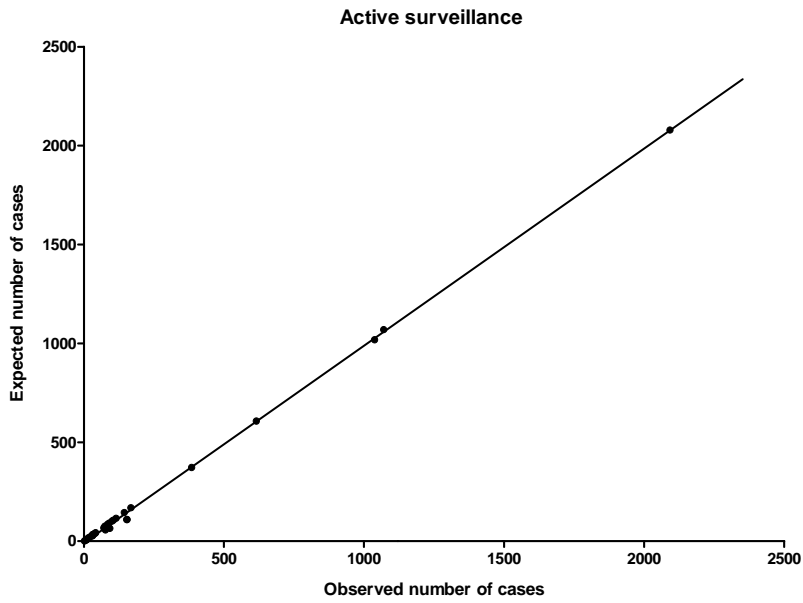


Figure 4: Plot of the observed and expected cases for the individual types for the active and the passive surveillance
 For Enteritidis and Typhimurium subtypes, the deterministically reallocated numbers of cases are used in place of the observed numbers of cases

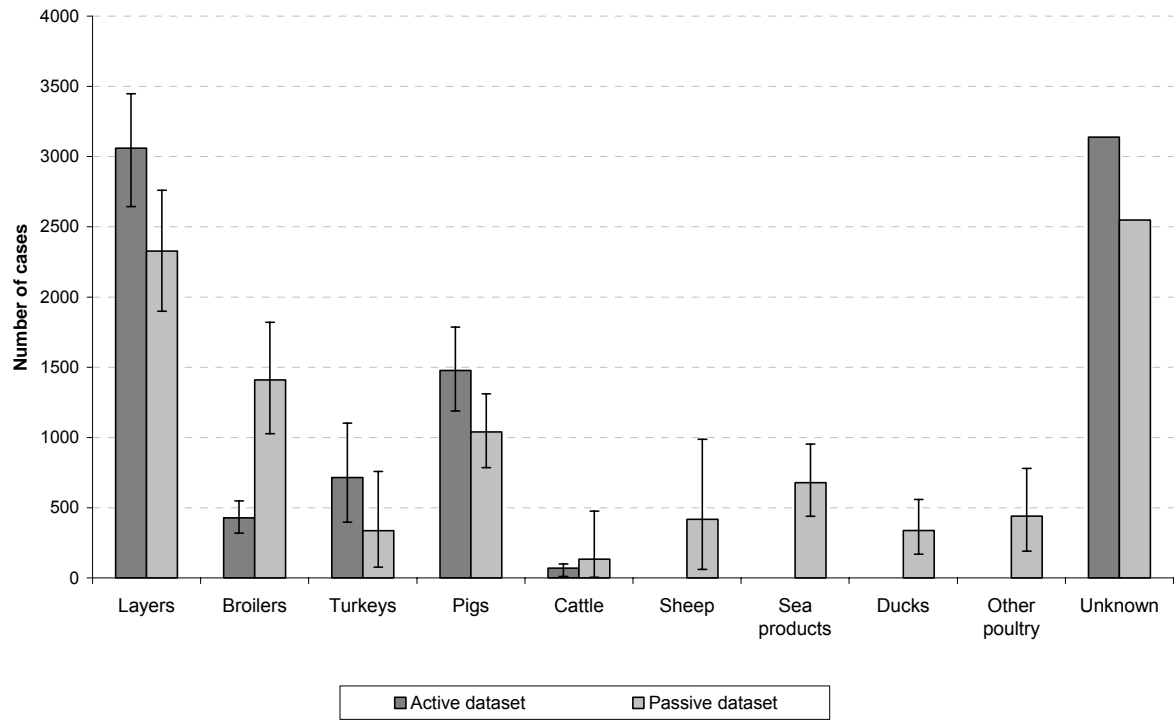


Figure 5: Posterior means and 95% CI of attributed cases per source according to surveillance mode

Discussion et conclusion générales

Résumé des résultats

Nos objectifs pour ces travaux étaient

- d'évaluer la possibilité de créer, à partir des données de surveillance disponibles en France, une base de données autorisant l'usage de la démarche d'attribution par typage microbiologique ;
- puis d'évaluer la robustesse du modèle et son adaptabilité à des données d'une qualité différente de celles pour lesquelles il avait été mis au point.

Grâce à un système de surveillance intégré basé sur une longue expérience et développé dans le cadre réglementaire européen, l'ensemble des données nécessaires a pu être centralisé. Cependant, celles-ci n'étaient pas directement utilisables dans le cadre d'une démarche d'attribution et ont nécessité un travail d'adaptation important, notamment en ce qui concerne le sous-typage. Ce travail a été rendu possible par l'implication de l'ensemble des acteurs concernés. Ainsi, un jeu de données suffisant pour appliquer la méthode de façon efficace a été constitué. Cependant, ses limites et particulièrement la non prise en compte de certaines sources potentielles, rendent l'interprétation des résultats produits hasardeuse.

Il a ensuite été montré que le modèle est particulièrement sensible à l'information informative choisie pour les priors. Déterminer les paramètres sur lesquels cette information informative doit porter (types spécifiques à une source) ainsi que les valeurs associées, en fonction des données a permis d'obtenir des résultats plus cohérents que l'option originale basée sur des choix arbitraires. Cette solution permet ainsi d'adapter le modèle à tout jeu de données, à condition que celui-ci contienne un nombre suffisant de types spécifiques.

Enfin, le jeu de données français nous a permis d'aborder le problème de l'impact de la qualité des données. Nous disposons en effet, pour ce qui concerne la contamination des sources, à la fois de données de surveillance active telles que préconisées par les développeurs du modèle, et de données de surveillance passive, considérées comme moins bien adaptées à ce type de démarche. Les données de surveillance passive ont permis d'obtenir des résultats d'attribution cohérents avec ceux obtenus avec les données

recommandées (surveillance active) tout en permettant d'inclure un plus grand nombre de sources, malgré une différence significative pour une filière. Enfin, ces travaux ont révélé l'intérêt d'une utilisation du modèle en parallèle sur les deux types de données. Ceci permet potentiellement de détecter des sources d'importance non surveillées activement et ainsi éventuellement d'orienter la surveillance active.

Les données : limites et opportunités

Le sous-typage des souches de Typhimurium et d'Enteritidis a été un point clé de ces travaux. Pour des raisons pragmatiques, la méthode choisie a été l'antibiotypie. Facile à mettre en œuvre, elle s'est révélée suffisamment discriminante et adaptée à une utilisation en attribution par typage microbiologique. Cependant, l'instabilité de certains traits de résistance (liés à des gènes situés sur des éléments génétiques mobiles) peut altérer la robustesse de cette méthode. De plus, si le modèle est utilisé pour attribuer les cas en fonction de la résistance aux antibiotiques des souches, l'antibiotypie ne pourra pas être utilisée pour le sous-typage. Pour des raisons de contraintes logistiques également, le nombre de souches humaines d'Enteritidis et de Typhimurium sous-typées était réduit (moins de 4% des cas de ces sérotypes). Le faible nombre de souches testées en électrophorèse en champ pulsé associé au faible pouvoir discriminant de cette méthode pour Enteritidis (quand elle est appliquée selon le protocole PulseNet avec une seule enzyme de restriction (Zheng, Keys et al. 2007)) explique probablement l'échec de l'utilisation de cette méthode de sous-typage pour l'utilisation en attribution. De plus, le faible nombre de souches humaines testées en antibiotypie crée de l'incertitude dans le modèle. En effet, il suppose la mise en place d'une étape de réallocation permettant d'extrapoler la distribution de sous-types observée parmi les souches testées à l'ensemble des cas d'Enteritidis et de Typhimurium. L'incertitude sur le nombre de cas réalloués utilisé dans le modèle sera d'autant plus grande que le nombre de souches testées est faible.

La méthode de sous-typage est donc un point clé pour une utilisation optimale du modèle. Dans cette optique, une réflexion commune des laboratoires responsables du typage des souches (Centre National de Référence et Laboratoire National de Référence des salmonelles) permettrait d'envisager l'adoption d'une méthode typage adaptée (discriminante pour Enteritidis et Typhimurium), standardisée, et dont le coût serait suffisamment réduit pour être utilisée en routine. La méthode de typage génotypique et automatisable, récemment brevetée par le Centre National de Référence des salmonelles, pourrait être une

bonne candidate. Cette méthode (méthode CRISPR _Clustered Regularly Interspaced Short Palindromic Repeats_), est basée sur le séquençage de deux loci de type CRISPR dans le génome de *Salmonella* et met en jeu l'utilisation d'une puce à ADN.

Lors de l'introduction de nouvelles sources dans le modèle, les cas attribués à celles-ci sont en partie des cas précédemment attribués aux sources initialement incluses dans le modèle et en partie des cas nouvellement attribués, qui étaient précédemment de source inconnue. Ainsi, augmenter le nombre de sources dans le modèle permet d'éviter les biais (cas attribués à tort), d'augmenter le nombre global de cas attribués et concomitamment de diminuer la proportion (très forte pour les données françaises) de cas d'origine inconnue. Le système de surveillance passive permet de répondre à cet enjeu. Il comprend en effet des données concernant un grand nombre de sources : différents animaux de rentes, animaux de compagnie et produits d'origine végétale. Cependant, pour ce qui concerne les denrées d'origine importée, il n'existe pas à l'heure actuelle de données suffisantes (sous-type notamment) dans cette base de données ni dans aucune autre à notre connaissance. Dans l'optique d'utiliser les résultats du modèle pour mettre en place des interventions en élevage, il est nécessaire de pouvoir distinguer la part réelle de cas liée aux productions nationales de celle liée aux produits importés. Dans le cadre de la Food Law, la Direction Générale de l'Alimentation contrôle les denrées importées issues de volailles. Les données ainsi collectées, pourraient être utilisées, sous condition de sous-typer les souches prélevées. L'idéal serait de pouvoir étendre ces contrôles à d'autres filières. D'autre part, il est probable qu'il existe dans la base de données passive du Réseau Salmonella des données concernant des produits importés. Inclure l'origine des échantillons (produit national ou importé) dans les fiches de recueil permettrait d'utiliser ces informations.

Enfin, la base de données INCA, utilisée pour extraire les données de consommation permet d'avoir accès non seulement aux quantités réellement consommées, mais également aux fréquences de consommation, le tout par classe d'âge. L'utilisation de la fréquence de consommation au lieu des quantités consommées, permettrait ainsi d'être plus en accord avec l'hypothèse d'exposition poissonnienne de la population. Elle permettrait de plus, de tenir compte du fait que, les portions standards étant différentes d'un aliment à l'autre, le lien entre quantités consommées et fréquence d'exposition au danger n'est pas de type linéaire. D'autre part, il apparaît dans les résultats de l'enquête INCA, que l'exposition des enfants est différente de celle des adultes. Comme le taux d'infection est également différent entre ces deux populations, il serait sans doute pertinent de les considérer séparément dans le modèle.

Le modèle : limites, adaptabilité

Différentes hypothèses sont nécessaires à la mise place du modèle. Tout d'abord, celui-ci s'applique aux cas domestiques, considérant que les cas ayant récemment voyagé ont été infectés à l'étranger. Dans notre jeu de données, l'item concernant le voyage était rarement renseigné. D'autre part, seuls les cas sporadiques sont inclus, ainsi qu'un cas par foyer. Ce parti pris peut être discuté. D'une part, comme le soulignent Greig et Ravel (2009), l'ampleur d'un foyer épidémique tient autant au mode de production et au mode de consommation de l'aliment incriminé qu'à l'aliment en lui-même, ce qui justifie de ne les compter que pour un cas. De plus, Hald *et al*/soulignent le risque de surestimer l'impact des aliments à l'origine de TIAC. D'autre part, inclure l'ensemble des cas sans distinction sporadique/cas groupés, pourrait permettre de « nourrir » le facteur source dépendant supposé mesurer la capacité de la source à véhiculer les salmonelles, à conditions que les cas de TIAC en question soient correctement attribués. La majorité des TIAC étant liée à Enteritidis et Typhimurium, cette deuxième option doit donc s'accompagner d'un sous-typage efficace et généralisé. Enfin, il est vraisemblable que, justement en raison de l'absence de données de sous-typage concernant les TIAC, seule une petite partie de celles-ci aient été détectées dans la base de données du CNR. Les cas humains considérés contiennent donc à l'évidence déjà une part de cas groupés. Enfin, il est supposé dans le modèle que l'ensemble des cas inclus sont liés aux sources considérées, ce qui semble peu réaliste au vu de la redistribution des cas, observée lors de l'inclusion de nouvelles sources dans le modèle.

Les données incluses dans le modèle doivent répondre à différents critères de qualité et pré-requis. Pour ce qui concerne la contamination des sources, les concepteurs du modèle recommandent d'utiliser des données de prévalence et insistent sur leur représentativité. Les travaux présentés au chapitre V, montrent que des proportions utilisées en remplacement des prévalences donnent des résultats satisfaisants. De même, l'utilisation de données passives dont la représentativité n'est pas assurée a donné des résultats cohérents avec ceux obtenus à partir de données conformes aux exigences exprimées ci-dessus. Ces exigences qualitatives concernant les données sont donc à nuancer. Hald *et al*/insistent par ailleurs sur la nécessité d'appliquer une méthode de sous-typage efficace, ce qui est effectivement indispensable (Cf § précédent). Enfin, le pré-requis d'une distribution hétérogène des types parmi les sources est incontournable et doit être vérifié.

De par sa structure, le modèle proposé par Hald *et al*/nécessite l'introduction d'information *a priori* informative, ce qui, dans un cadre bayésien, est inadapté. Cependant, cette

information informative est indispensable en raison de la surparamétrisation du modèle. Dans ce cadre, baser le choix des paramètres à fixer (parmi les paramètres spécifiques aux sources et les paramètres spécifiques aux types) et des valeurs correspondantes, sur les données permet de limiter l'impact de l'information sur le modèle. En revanche, cela conduit à une surexploitation de l'information contenue dans les données. Une meilleure solution serait de disposer de mesures répétées ou de définir les paramètres source et type dépendants afin de pouvoir les « nourrir » à partir d'information exogène (données d'analyse quantitative de risque, pathogénicité, virulence, ...).

Notre proposition de paramétrisation porte sur les paramètres type dépendants. Les paramètres fixés correspondent aux types spécifiques qui sont définis, en fonction des observations, comme des types microbiologiques présents dans une seule filière. Cette définition dépend de plusieurs éléments : typage utilisé, prévalence du type considéré, taille de l'échantillon testé (limite de détection), fitness des souches, nombre de sources considérées. Ainsi, dans les deux jeux de données dont nous disposons, peu de sérotypes spécifiques étaient communs. La notion de type spécifique serait à explorer plus avant. On pourrait notamment considérer la possibilité de définir les types spécifiques de façon plus universelle à partir de données épidémiologiques plus larges et de connaissances biologiques sur l'écologie des sérotypes (le sérotype S 61:k:1,5,7, par exemple, est également fortement associé au mouton en Norvège (Alvseike and Skjerve 2002)), tout en tenant compte des limites de l'exercice, liées à l'évolution rapide de la distribution des sérotypes et à l'origine des souches (aliments contaminés distribués à grande échelle par exemple). Enfin, cette paramétrisation rajoute un pré-requis au modèle, elle ne pourra être utilisée que si suffisamment de types spécifiques sont présents dans le jeu de données.

L'équipe de Tine Hald a récemment présenté, dans le cadre du groupe de travail européen, une approche d'attribution intégrant la notion de temps. L'intégration d'une troisième dimension (en plus du type et de la source) rend la structure du modèle plus complexe (la question de la sensibilité du modèle n'a toutefois pas été discutée). En revanche, cela permet de disposer de mesures répétées. Les conclusions qui seront tirées de ces travaux seront donc à considérer avec intérêt.

Ainsi, nous avons mis en lumière la sensibilité attendue du modèle à l'information *a priori* informative ainsi que les qualités des données indispensables (méthode de sous-typage efficace) ou supportant un écart (prévalence, représentativité) pour une utilisation dans le modèle. Il serait cependant nécessaire d'aller plus loin : analyse des conditions limites d'utilisation, analyse de sensibilité globale à partir de données simulées, définition des

facteurs source et type dépendants. Pour répondre à ces objectifs, un post-doctorat a d'ores et déjà été sollicité dans le prolongement de ces travaux de thèse à l'AFSSA dans l'équipe du Pôle d'Appui Scientifique à l'Evaluation des Risques (Maisons-Alfort).

Intérêt de la méthode d'attribution par typage microbiologique

Les résultats d'attribution obtenus (chapitre V) sont similaires en terme d'ordre d'importance des sources à ceux obtenus au Danemark (chapitre III). Pour les résultats obtenus avec les données de surveillance active, ils reflètent finalement, pour les filières majoritairement consommées en France, les données de prévalence globale par source. Il serait intéressant d'étudier les résultats obtenus pour une filière à forte prévalence mais faible consommation. De plus, la forte proportion de cas de source inconnue interpelle. L'amélioration du sous-typage d'une part, et l'inclusion de sources non prises en compte dans le modèle d'autre part devraient permettre de limiter le nombre de cas non attribuables. Enfin, l'outil d'attribution permet, en l'absence de données de prévalence, de hiérarchiser l'importance des sources surveillées de façon passive.

Au-delà des résultats obtenus avec le jeu de données français, avec les défauts que celui-ci comporte, la démarche d'attribution est une approche de première importance. Il s'agit de la première approche intégrative verticale et horizontale qui permet d'intégrer les données concernant les cas, les sources, le niveau d'exposition (consommation) et ce globalement pour plusieurs sources. L'approche d'attribution par typage microbiologique en particulier permet de faire le lien entre la contamination de différents types d'élevages et leur impact respectif sur la santé humaine. Elle offre donc l'opportunité de pouvoir évaluer l'impact d'interventions qui concernent parfois plusieurs filières en même temps, ou plus simplement de déterminer l'impact propre d'une intervention sur l'évolution des infections humaines, dans un contexte où la contribution relative des différentes sources évolue très probablement dans le temps. Cette démarche a pu être appliquée en France grâce au système de surveillance intégré dont la partie concernant le secteur agro-vétérinaire est bien développé, contrairement à des pays comme les Etats-Unis. Si les conditions d'utilisation ne sont pas encore optimales (absence de données sur les denrées importées par exemple), la dynamique des collaborations qui se sont construites autour de ce projet, devrait permettre de progresser. De plus, ce projet a permis de mettre en lumière la complémentarité entre les systèmes de surveillance active et passive pour ce qui concerne la contamination des

sources, et l'utilisation potentielle des données passives dans le modèle pour piloter la surveillance active, en détectant les filières émergentes en terme d'impact sur les cas humains.

Dans l'optique d'une utilisation du modèle en routine, pour évaluer l'impact d'interventions par exemple, il serait nécessaire de collecter les données de surveillance active de façon plus régulière (annuelle si possible), ce qui devrait bientôt être le cas suite à l'implémentation de la directive zoonoses. De plus, l'évolution annuelle de l'ordre d'importance des sérotypes et des profils de résistance, rendra nécessaire de redéfinir à chaque fois les sérotypes à inclure dans le modèle ainsi que les sous-types d'Enteritidis et Typhimurium. Si la comparabilité des résultats d'une année à l'autre sera à discuter, les règles d'inclusion des sérotypes et la méthode statistique de formation des clusters de profils d'antibiorésistance, ont été pensées pour être répétables. L'utilisation de ce type d'approche est actuellement à l'étude au niveau européen. Les données de surveillance obligatoire des filières ainsi que celles concernant les cas humains sont en effet centralisées au niveau de l'EFSA (European Food Safety Agency) et de l'ECDC (European Center for Disease Control) et une réflexion est menée sur les approches d'attribution. Il est donc d'autant plus essentiel pour la France d'avoir un regard critique sur ces méthodes et ce qui peut en être attendu.

Une perspective majeure de l'approche, est son utilisation pour attribuer les cas en fonction de leur résistance aux antibiotiques. L'importance de la problématique antibiorésistance en santé publique n'est plus à démontrer, et l'impact de l'utilisation des antibiotiques en élevage sur la santé humaine est au centre de nombreux travaux. Le modèle d'attribution par typage microbiologique, qui a déjà été utilisé dans ce sens (Hald, Lo Fo Wong et al. 2007), devrait permettre de faire le lien entre les modes d'utilisation des antibiotiques, parfois très différents d'une filière animale à l'autre, et les cas humains rapportés.

Enfin, différentes approches d'attribution existent qui permettent de considérer des points d'attribution autres que le réservoir et ainsi d'ouvrir des perspectives pour l'identification d'autres types d'intervention. La méthode d'évaluation comparative de l'exposition, par exemple, permet de considérer différentes voies de contamination comme le contact avec les animaux de compagnie, les contaminations alimentaires ou encore les eaux de baignades, ce qui peut être précieux pour établir les cibles de potentielles campagnes d'éducation de la population. A notre connaissance, aucune de ces méthodes n'a été appliquée en France à ce jour. Il serait pourtant intéressant de pouvoir comparer les résultats d'attribution obtenus ici avec différentes méthodes et ainsi de valoriser les riches données de surveillance françaises.

Conclusion

Ce travail de thèse a permis d'identifier les limites du modèle concernant la démarche proposée et les données pouvant être utilisées. Les atouts actuels du dispositif national de surveillance ont été mis en avant et des améliorations méthodologiques ainsi que des perspectives de développement ont été proposées tant dans le domaine de la surveillance épidémiologique des salmonelloses humaines et des salmonelles chez les productions animales que dans l'évaluation du risque. L'amélioration continue de cette démarche suppose un dialogue permanent entre les évaluateurs et les gestionnaires du risque pour une utilisation optimale en santé publique, pour le ciblage des études de surveillance active et l'amélioration des données de surveillance passive d'une part, comme pour le ciblage des interventions et la mesure de leurs effets sur le moyen terme d'autre part.

Références

- Allard, D. G. (2002). "The 'farm to plate' approach to food safety - Everyone's business." Canadian Journal of Infectious Diseases **13**(3): 185-190.
- Allos, B. M., M. R. Moore, et al. (2004). "Surveillance for sporadic foodborne disease in the 21st century: the FoodNet perspective." Clin Infect Dis **38 Suppl 3**: S115-20.
- Alvseike, O. and E. Skjerve (2002). "Prevalence of a *Salmonella* subspecies *diarizonae* in Norwegian sheep herds." Preventive Veterinary Medicine **52**(3-4): 277-285.
- Ammon, A. and R. V. Tauxe (2007). "Investigation of multi-national foodborne outbreaks in Europe: some challenges remain." Epidemiol Infect **135**(6): 887-9.
- Angulo, F. J., K. R. Johnson, et al. (2000). "Origins and consequences of antimicrobial-resistant nontyphoidal *Salmonella*: Implications for the use of fluoroquinolones in food animals." Microbial Drug Resistance **6**(1): 77-83.
- Angulo, F. J., A. C. Voetsch, et al. (1998). "Determining the burden of human illness from food borne diseases. CDC's emerging infectious disease program Food Borne Diseases Active Surveillance Network (FoodNet)." The Veterinary clinics of North America. Food animal practice **14**(1): 165-172.
- Anonymous (2001). Rapport au parlement européen et au conseil sur les mesures à mettre en oeuvre pour le contrôle et la prévention des zoonoses, Commission des communautés européennes: 78.
- Anonymous (2005). "Reported foodborne illness and gastroenteritis in Australia: annual report of the OzfoodNet network, 2004." Communicable diseases intelligence **29**(2): 165-192.
- Anonymous (2006). "Report of Task Force on Zoonoses Data Collection on Guidance Document on Good Practices for Design of Field Surveys." The EFSA Journal **93**: 1-24.
- Anonymous (2009a). "The community summary report on trends and sources of zoonoses and zoonotic agents in the european union in 2007." The EFSA Journal **223**: 217.
- Anonymous (2009b). The Rapid Alert System for Food and Feed (RASFF), Annual Report 2008, Directorate-General for Health and Consumers, European Commission: 56.
- Arsenault, J., A. Letellier, et al. (2007c). "Prevalence and risk factors for *Salmonella* and *Campylobacter* spp. carcass contamination in broiler chickens slaughtered in Quebec, Canada." J Food Prot **70**(8): 1820-8.
- Arsenault, J., A. Letellier, et al. (2007b). "Prevalence and risk factors for *Salmonella* and *Campylobacter* spp. carcass contamination in turkeys slaughtered in Quebec, Canada." J Food Prot **70**(6): 1350-9.
- Bager, F. and R. Helmuth (2001). "Epidemiology of resistance to quinolones in *Salmonella*." Vet Res **32**(3-4): 285-90.
- Barber, D. A., G. Y. Miller, et al. (2003). "Models of antimicrobial resistance and foodborne illness: examining assumptions and practical applications." J Food Prot **66**(4): 700-9.
- Baroukh, T., Y. Le Strat, et al. (2008). Use of statistical methods for a routinely unusual event surveillance in non human *Salmonella*. European Scientific Conference on Applied Infectious Disease Epidemiology (ESCAIDE), Berlin.
- Barza, M. (2002). "Potential mechanisms of increased disease in humans from antimicrobial resistance in food animals." Clin Infect Dis **34 Suppl 3**: S123-5.
- Batz, M. B., M. P. Doyle, et al. (2005). "Attributing illness to food." Emerging Infectious Diseases **11**(7): 993-999.

- Bellido Blasco, J. B., J. M. Gonzalez Cano, et al. (1998). "Factors associated with sporadic cases of salmonellosis in 1- to 7-year-old children. Study of cases and controls." Factores asociados con casos esporádicos de salmonelosis en niños de 1 a 7 años. Estudio de casos y controles. *Gaceta sanitaria / S.E.S.P.A.S* **12**(3): 118-125.
- Beloëil, P. A. (2007). Epidémiologie analytique de *Salmonella enterica* et *Listeria Monocytogenes* en production primaire porcine (élevages hors-sol de type naisseur-engraisseur). Sciences biologiques et médicales. Bordeaux, Université de Bordeaux 2: 355.
- Bender, J. B., C. W. Hedberg, et al. (2001). "Use of molecular subtyping in surveillance for salmonella enterica serotype typhimurium." New England Journal of Medicine **344**(3): 189-195.
- Berge, A. C. B., E. R. Atwill, et al. (2003). "Assessing antibiotic resistance in fecal *Escherichia coli* in young calves using cluster analysis techniques." Preventive Veterinary Medicine **61**(2): 91-102.
- Binkowitz, B. S. and D. Wartenberg (2001). "Disparity in quantitative risk assessment: a review of input distributions." Risk Anal **21**(1): 75-90.
- Blaser, M. J. and L. S. Newman (1982). "A review of human salmonellosis: I. Infective dose." Reviews of Infectious Diseases **4**(6): 1096-1106.
- Bollaerts, K., M. Aerts, et al. (2008). "Human salmonellosis: estimation of dose-illness from outbreak data." Risk Anal **28**(2): 427-40.
- Botrel, M. A., M. Chazel, et al. (2006). "Le RESAPATH : Analyse critique et propositions d'amélioration." Epidémiologie et santé animale **50**: 157-168.
- Botteldoorn, N., L. Herman, et al. (2004). "Phenotypic and molecular typing of *Salmonella* strains reveals different contamination sources in two commercial pig slaughterhouses." Applied and Environmental Microbiology **70**(9): 5305-5314.
- Boxrud, D., K. Pederson-Gulrud, et al. (2007). "Comparison of multiple-locus variable-number tandem repeat analysis, pulsed-field gel electrophoresis, and phage typing for subtype analysis of *Salmonella enterica* serotype Enteritidis." J Clin Microbiol **45**(2): 536-43.
- Brandl, M. T. (2006). "Fitness of human enteric pathogens on plants and implications for food safety." Annu Rev Phytopathol **44**: 367-92.
- Brisabois, A. (2001). "Intérêt et limites des techniques de caractérisation des *Salmonella*." Epidémiologie et santé animale **39**: 31-42.
- Brooks, S. P. and A. Gelman (1998). "General methods for monitoring convergence of iterative simulations." Journal of Computational and Graphical Statistics **7**(4): 434-455.
- Brooks, S. P. and G. O. Roberts (1998b). "Convergence assessment techniques for Markov chain Monte Carlo." Statistics and Computing **8**(4): 319-335.
- Brouard, C., E. Espié, et al. (2007). "Two consecutive large outbreaks of *Salmonella enterica* serotype Agona infections in infants linked to the consumption of powdered infant formula." Pediatr Infect Dis J **26**(2): 148-52.
- Butt, A. A., K. E. Aldridge, et al. (2004). "Infections related to the ingestion of seafood Part I: Viral and bacterial infections." Lancet Infectious Diseases **4**(4): 201-212.
- CDC (2001). "Updated guidelines for evaluating public health surveillance systems." MMWR **50**(RR13): 1-35.
- Cezard, J. P., J. P. Chouraqui, et al. (2002). "[Drug treatment of acute infectious diarrhea in infants and children]." Arch Pediatr **9**(6): 620-8.
- Chazel, M., Y. Buret, et al. (2007). "Le RESSAB - Réseau d'Epidémiologie et de Surveillance des Salmonelloses Bovines - Résultats 2006." Bulletin Epidémiologique - AFSSA **25**: 5-7.
- Cloëckaert, A. and E. Chaslus-Dancla (2001). "Mechanisms of quinolone resistance in *Salmonella*." Vet Res **32**(3-4): 291-300.
- Coleman, M. E., H. M. Marks, et al. (2004). "Discerning strain effects in microbial dose-response data." J Toxicol Environ Health A **67**(8-10): 667-85.

- Combris, P. (1997). "The consumption of animal products in France : current tendencies and future perspectives." Productions Animales **10**(4): 267-274.
- Cowles, M. K. and B. P. Carlin (1996). "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review." Journal of the American Statistical Association **91**(434): 883-904.
- D'Aoust, J.-Y. (1991). "Pathogenicity of foodborne Salmonella." International Journal of Food Microbiology **12**(1): 17-40.
- D'Aoust, J. (1989). Salmonella. Foodborne Bacterial Pathogens. M. P. Doyle. New York, Marcel Dekker: 327-445.
- D'Aoust, J. (1994). "Salmonella and the international food trade." International Journal of Food Microbiology **24**(1-2): 11-31.
- David, J., C. Danan, et al. (2008). Developing a model to quantify the contribution of the major food-animal sources to human cases of salmonellosis regarding their sensibility to antimicrobials in France. 4th Med-Vet-Net annual scientific meeting, Saint Malo, France.
- David, J. M., C. Danan, et al. (submitted). "Structure of the French farm to table surveillance system for *Salmonella*."
- De Jong, B., Y. Andersson, et al. (2005). "Effect of regulation and education on reptile-associated salmonellosis." Emerging Infectious Diseases **11**(3): 398-403.
- de Jong, B. and K. Ekdahl (2006b). "Human salmonellosis in travellers is highly correlated to the prevalence of salmonella in laying hen flocks." Euro Surveill **11**(7): E060706 1.
- Dechet, A. M., E. Scallan, et al. (2006). "Outbreak of multidrug-resistant Salmonella enterica serotype Typhimurium Definitive Type 104 infection linked to commercial ground beef, northeastern United States, 2003-2004." Clin Infect Dis **42**(6): 747-52.
- Delavigne, A. E. (2001). "Aliments purs, aliments sain(t)s, la dimension identitaire des peurs alimentaires au Danemark." Anthropology of food(0).
- Delgado Ronda, N., J. L. Munoz Bellido, et al. (2006). "Molecular epidemiology of drug-resistant Salmonella Typhimurium in Spain." Revista Espanola de Quimioterapia **19**(2): 152-160.
- Delmas, G., A. Gallay, et al. (2007). "Food borne infection and intoxication occurrence in France between 1996 and 2005." Les toxi-infections alimentaires collectives en France entre 1996 et 2005 **48**(278): 74-78.
- Denagamage, T. N., A. M. O'Connor, et al. (2007). "Efficacy of vaccination to reduce Salmonella prevalence in live and slaughtered swine: A systematic review of literature from 1979 to 2007." Foodborne Pathogens and Disease **4**(4): 539-549.
- Deplano, A., M. Struelens, et al. (2007). Marqueurs épidémiologiques. Précis de bactériologie Clinique. E. édition. Paris, Freney, J., Renaud, F., Leclercq, R., Riegel, P.: 1764.
- Desenclos, J. C., V. Vaillant, et al. (2002). "[Food-borne infections: do we need further data collection for public health action and research]." Rev Epidemiol Sante Publique **50**(1): 67-79.
- Devasia, R. A., J. K. Varma, et al. (2005). "Antimicrobial use and outcomes in patients with multidrug-resistant and pansusceptible Salmonella newport infections, 2002-2003." Microbial Drug Resistance **11**(4): 371-377.
- Dominguez, M., N. Jourdan-Da Silva, et al. (2008). "Outbreak of Salmonella enterica Serotype Montevideo Infections in France Linked to Consumption of Cheese Made from Raw Milk." Foodborne Pathog Dis.
- Doran, G., M. NiChulain, et al. (2006). "Interpreting streptomycin susceptibility test results for Salmonella enterica serovar Typhimurium." International Journal of Antimicrobial Agents **27**(6): 538-540.
- Dufour, B. and S. La Vieille (2000). "Epidemiological surveillance of infectious diseases in France." Vet Res **31**(2): 169-85.
- DuPont, H. L. (2007). "The growing threat of foodborne bacterial enteropathogens of animal origin." Clin Infect Dis **45**(10): 1353-61.

- Edel, W. (1994). "Salmonella enteritidis eradication programme in poultry breeder flocks in The Netherlands." International Journal of Food Microbiology **21**(1-2): 171-178.
- Ellerbroek, L., T. Alter, et al. (2009). "[Tasks and duties of veterinary reference laboratories for food borne zoonoses]." Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz **52**(2): 208-13.
- Elviss, N. C., C. L. Little, et al. (2009). "Microbiological study of fresh herbs from retail premises uncovers an international outbreak of salmonellosis." Int J Food Microbiol.
- Espié, E., H. De Valk, et al. (2005b). "An outbreak of multidrug-resistant Salmonella enterica serotype Newport infections linked to the consumption of imported horse meat in France." Epidemiol Infect **133**(2): 373-6.
- Espié, E., F. X. Weill, et al. (2005a). "Nationwide outbreak of Salmonella enterica serotype Agona infections in infants in France, linked to infant milk formula, investigations ongoing." Euro Surveill **10**(3): E050310 1.
- Evers, E. G., H. J. Van Der Fels-Klerx, et al. (2008). "Campylobacter source attribution by exposure assessment." International Journal of Risk Assessment and Management **8**(1-2): 174-190.
- Fakhr, M. K., L. K. Nolan, et al. (2005). "Multilocus sequence typing lacks the discriminatory ability of pulsed-field gel electrophoresis for typing Salmonella enterica serovar Typhimurium." Journal of Clinical Microbiology **43**(5): 2215-2219.
- Finley, R., R. Reid-Smith, et al. (2006). "Human health implications of Salmonella-contaminated natural pet treats and raw pet food." Clin Infect Dis **42**(5): 686-91.
- Flint, J. A., Y. T. Van Duynhoven, et al. (2005). "Estimating the burden of acute gastroenteritis, foodborne disease, and pathogens commonly transmitted by food: An international review." Clinical Infectious Diseases **41**(5): 698-704.
- Foley, S. L. and A. M. Lynne (2008). "Food animal-associated Salmonella challenges: pathogenicity and antimicrobial resistance." Journal of animal science **86**(14 Suppl).
- Foley, S. L., A. M. Lynne, et al. (2009). "Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens." Infection, Genetics and Evolution **9**(4): 430-440.
- Foxman, B., L. Zhang, et al. (2005). "Choosing an appropriate bacterial typing technique for epidemiologic studies." Epidemiologic Perspectives and Innovations **2**.
- Frenzen, P. D., T. L. Riggs, et al. (1999). "Salmonella Cost Estimate Updated Using FoodNet Data." FoodReview **22**(2): 10-15.
- Galanis, E., D. M. A. Lo Fo Wong, et al. (2006). "Web-based surveillance and global Salmonella distribution, 2000-2002." Emerging Infectious Diseases **12**(3): 381-388.
- Gallay, A., V. Vaillant, et al. (2000). "How many foodborne outbreaks of Salmonella infection occurred in France in 1995?" American Journal of Epidemiology **152**(2): 171-177.
- Gebreyes, W. A., C. Altier, et al. (2006). "Molecular epidemiology and diversity of Salmonella serovar typhimurium in pigs using phenotypic and genotypic approaches." Epidemiology and Infection **134**(1): 187-198.
- Gerner-Smidt, P., K. Hise, et al. (2006). "PulseNet USA: A five-year update." Foodborne Pathogens and Disease **3**(1): 9-19.
- Gervelmeyer, A., M. Hempen, et al. (2008). "Developing the Community reporting system for foodborne outbreaks." Euro Surveill **13**(45): pii: 19029.
- Gilks, W. R., S. Richardson, et al. (1996). Markov Chain Monte Carlo in practice: Interdisciplinary Statistics. London, Chapman & Hall/CRC.
- Gillespie, I. and R. Elson (2005). "Successful reduction of human Salmonella Enteritidis infection in England and Wales." Euro surveillance : bulletin europ en sur les maladies transmissibles = European communicable disease bulletin. **10**(11).
- Glynn, M. K., V. Reddy, et al. (2004). "Prior antimicrobial agent use increases the risk of sporadic infections with multidrug-resistant Salmonella enterica serotype typhimurium: A foodnet case-control study, 1996-1997." Clinical Infectious Diseases **38**(SUPPL. 3).

- Greig, J. D. and A. Ravel (2009). "Analysis of foodborne outbreak data reported internationally for source attribution." International Journal of Food Microbiology **130**(2): 77-87.
- Gudmundsdottir, S., H. Hardardottir, et al. (2003). "Subtyping of Salmonella enterica serovar Typhimurium outbreak strains isolated from humans and animals in Iceland." Journal of Clinical Microbiology **41**(10): 4833-4835.
- Guerra, B., P. Schrors, et al. (2000). "Application of PFGE performed with XbaI to an epidemiological and phylogenetic study of Salmonella serotype typhimurium. Relations between genetic types and phage types." New Microbiologica **23**(1): 11-20.
- Haeghebaert, S., P. Sulem, et al. (2003). "Two outbreaks of Salmonella enteritidis phage type 8 linked to the consumption of Cantal cheese made with raw milk, France, 2001." Euro Surveill **8**(7): 151-6.
- Hald, T., D. M. A. Lo Fo Wong, et al. (2007). "The attribution of human infections with antimicrobial resistant Salmonella bacteria in Denmark to sources of animal origin." Foodborne Pathogens and Disease **4**(3): 313-326.
- Hald, T., D. Vose, et al. (2004). "A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis." Risk Anal **24**(1): 255-69.
- Harbottle, H., D. G. White, et al. (2006). "Comparison of multilocus sequence typing, pulsed-field gel electrophoresis, and antimicrobial susceptibility typing for characterization of Salmonella enterica serotype Newport isolates." J Clin Microbiol **44**(7): 2449-57.
- Havelaar, A. H., J. Bräunig, et al. (2007). "Towards an integrated approach in supporting microbiological food safety decisions." Zoonoses and Public Health **54**(3-4): 103-117.
- Helms, M., J. Simonsen, et al. (2004). "Quinolone resistance is associated with increased risk of invasive illness or death during infection with Salmonella serotype Typhimurium." J Infect Dis **190**(9): 1652-4.
- Helms, M., J. Simonsen, et al. (2006). "Foodborne bacterial infection and hospitalization: a registry-based study." Clin Infect Dis **42**(4): 498-506.
- Helms, M., P. Vastrup, et al. (2002). "Excess mortality associated with antimicrobial drug-resistant Salmonella typhimurium." Emerg Infect Dis **8**(5): 490-5.
- Helms, M., P. Vastrup, et al. (2003). "Short and long term mortality associated with foodborne bacterial gastrointestinal infections: registry based study." Bmj **326**(7385): 357.
- Hendrikx, P., M. Chazel, et al. (2005). "Development of performance indicators for the bovine clinical salmonellosis surveillance network in France." J Vet Med B Infect Dis Vet Public Health **52**(10): 465-75.
- Hennessy, T. W., L. H. Cheng, et al. (2004). "Egg consumption is the principal risk factor for sporadic Salmonella serotype heidelberg infections: A case-control study in foodnet sites." Clinical Infectious Diseases **38**(SUPPL. 3).
- Herida, M. (2008). "Evaluation du dispositif des maladies à déclaration obligatoire en France : connaissances, attitudes et pratiques des médecins et des biologistes, 2005." Bulletin épidémiologique hebdomadaire **51-52**: 503-507.
- Hohmann, E. L. (2001). "Nontyphoidal salmonellosis." Clinical Infectious Diseases **32**(2): 263-269.
- Hueston, W. D. (1993). "Assessment of national systems for the surveillance and monitoring of animal health." Rev Sci Tech **12**(4): 1187-96.
- Hyytia-Trees, E. K., K. Cooper, et al. (2007). "Recent developments and future prospects in subtyping of foodborne bacterial pathogens." Future Microbiology **2**(2): 175-185.
- ICMSF, I. C. o. M. S. f. F. (2006). "Use of epidemiologic data to measure the impact of food safety control programs." Food Control **17**(10): 825-837.
- InVS (2004). Morbidité et mortalité dues aux maladies infectieuses d'origine alimentaire en France: 192.
- Jeoffreys, N. J., G. S. James, et al. (2001). "Practical evaluation of molecular subtyping and phage typing in outbreaks of infection due to Salmonella enterica serotype Typhimurium." Pathology **33**(1): 66-72.

- Jones, T. F. and F. J. Angulo (2006). "Eating in restaurants: A risk factor for foodborne disease?" Clinical Infectious Diseases **43**(10): 1324-1328.
- Jones, T. F., L. A. Ingram, et al. (2008). "Salmonellosis outcomes differ substantially by serotype." J Infect Dis **198**(1): 109-14.
- Jourdan, N., S. Le Hello, et al. (2008). "Nationwide outbreak of Salmonella enterica serotype Give infections in infants in France, linked to infant milk formula, September 2008." Euro Surveill **13**(39).
- Jouy, E., D. Meunier, et al. (2003). "RESAPATH : Réseau de surveillance de la résistance aux antibiotiques chez les principales bactéries pathogènes des bovins, des porcs et des volailles. Description des espèces bactériennes et des pathologies enregistrées en 2002." Bulletin Epidémiologique - AFSSA **9**: 5-6.
- Kapperud, G., J. Lassen, et al. (1998). "Salmonella infections in Norway: Descriptive epidemiology and a case-control study." Epidemiology and Infection **121**(3): 569-577.
- Kariuki, S., C. Gilks, et al. (1999). "Analysis of Salmonella enterica serotype Typhimurium by phage typing, antimicrobial susceptibility and pulsed-field gel electrophoresis." Journal of Medical Microbiology **48**(11): 1037-1042.
- Kennedy, M., R. Villar, et al. (2004). "Hospitalizations and deaths due to Salmonella infections, FoodNet, 1996-1999." Clin Infect Dis **38 Suppl 3**: S142-8.
- Kerouanton, A., M. Marault, et al. (2007). "Pulsed-field gel electrophoresis subtyping database for foodborne Salmonella enterica serotype discrimination." Foodborne Pathog Dis **4**(3): 293-303.
- Kimura, A. C., V. Reddy, et al. (2004). "Chicken consumption is a newly identified risk factor for sporadic Salmonella enterica serotype enteritidis infections in the United States: A case-control study in FoodNet sites." Clinical Infectious Diseases **38**(SUPPL. 3).
- Kirk, M. D., I. McKay, et al. (2008). "Food safety: foodborne disease in Australia: the OzFoodNet experience." Clin Infect Dis **47**(3): 392-400.
- Kohl, K. S., K. Rietberg, et al. (2002). "Relationship between home food-handling practices and sporadic salmonellosis in adults in Louisiana, United States." Epidemiology and Infection **129**(2): 267-276.
- Korsak, N., Clinquart, A., Daube, G. (2004). "Salmonella spp. dans les denrées alimentaires d'origine animale : un réel problème de santé publique ?" Les annales de médecine vétérinaire **148**(4): 174-193.
- Lailier, R., M. Sanaa, et al. (2005). "Prevalence of multidrug resistant (MDR) Salmonella in bovine dairy herds in western France." Preventive Veterinary Medicine **70**(3/4): 177-189.
- Le Minor, L. and M. Y. Popoff (1987). "Designation of *Salmonella enterica* sp. nov. as the type and the only species of the genus *Salmonella*." International Journal of Systematic Bacteriology **37**: 465-468.
- Le Minor, L., M. Y. Popoff, et al. (1986). "Description of a seventh Salmonella subspecies: *S. choleraesuis* subsp. *indica* subsp. nov." Annales de microbiologie, Institut Pasteur **137 B**(2): 211-217.
- Lebart, L., A. Morineau, et al. (2000). Statistique exploratoire multidimensionnelle. Paris, Dunod.
- Liebana, E., L. Garcia-Migura, et al. (2002a). "Multiple genetic typing of Salmonella enterica serotype typhimurium isolates of different phage types (DT104, U302, DT204b, and DT49) from animals and humans in England, Wales, and Northern Ireland." Journal of Clinical Microbiology **40**(12): 4450-4456.
- Lo Fo Wong, D. M. A., T. Hald, et al. (2002). "Epidemiology and control measures for Salmonella in pigs and pork." Livestock Production Science **76**(3): 215-222.
- Lukinmaa, S., U.-M. Nakari, et al. (2006). "Genomic diversity within phage types of Salmonella enterica ssp. enterica serotypes enteritidis and typhimurium." Foodborne Pathogens and Disease **3**(1): 97-105.

- Lunn, D., D. Spiegelhalter, et al. (2009). "The BUGS project: Evolution, critique and future directions." Stat Med.(28): 3049-3067.
- Lunn, D. J., A. Thomas, et al. (2000). "WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility." Statistics and Computing **10**(4): 325-337.
- Malorny, B., A. Schroeter, et al. (2001). "Evaluation of molecular typing methods for Salmonella enterica serovar Typhimurium DT104 isolated in Germany from healthy pigs." Veterinary Research **32**(2): 119-129.
- Maraki, S., G. Samonis, et al. (2006). "Serotype distribution and antimicrobial resistance of Salmonella enterica from patients with gastroenteritis in Crete, Greece." Eur J Clin Microbiol Infect Dis **25**(2): 116-9.
- Marcus, R., J. K. Varma, et al. (2007). "Re-assessment of risk factors for sporadic Salmonella serotype Enteritidis infections: A case-control study in five FoodNet Sites, 2002-2003." Epidemiology and Infection **135**(1): 84-92.
- Martin, L. J., M. Fyfe, et al. (2004). "Increased Burden of Illness Associated with Antimicrobial-Resistant Salmonella enterica Serotype Typhimurium Infections." Journal of Infectious Diseases **189**(3): 377-384.
- Martinez, J. L. and F. Baquero (2002). "Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance." Clin Microbiol Rev **15**(4): 647-79.
- Mather, A. E., D. T. Haydon, et al. (2009). Comparing active and passive surveillance of antimicrobial resistance: Are we getting the whole story? 12th International Symposium on Veterinary Epidemiology and Economics, Durban.
- Mead, P. S., L. Slutsker, et al. (1999). "Food-related illness and death in the United States." Emerging Infectious Diseases **5**(5): 607-625.
- Mermin, J., L. Hutwagner, et al. (2004). "Reptiles, amphibians, and human Salmonella infection: A population-based, case-control study." Clinical Infectious Diseases **38**(SUPPL. 3).
- Millemann, Y., S. Gaubert, et al. (2000). "Evaluation of IS200-PCR and comparison with other molecular markers To trace Salmonella enterica subsp. enterica serotype typhimurium bovine isolates from farm to meat." J Clin Microbiol **38**(6): 2204-9.
- Miller, G. Y., X. Liu, et al. (2005). "Influence of Salmonella in pigs preharvest and during pork processing on human health costs and risks from pork." Journal of Food Protection **68**(9): 1788-1798.
- Miriagou, V., P. T. Tassios, et al. (2004). "Expanded-spectrum cephalosporin resistance in non-typhoid Salmonella." Int J Antimicrob Agents **23**(6): 547-55.
- Molbak, K. (2005). "Human health consequences of antimicrobial drug-resistant Salmonella and other foodborne pathogens." Clin Infect Dis **41**(11): 1613-20.
- Mølbak, K. (2004). "Spread of resistant bacteria and resistance genes from animals to humans - The public health consequences." Journal of Veterinary Medicine Series B: Infectious Diseases and Veterinary Public Health **51**(8-9): 364-369.
- Molbak, K., P. Gerner-Smidt, et al. (2002b). "Increasing quinolone resistance in Salmonella enterica serotype Enteritidis." Emerg Infect Dis **8**(5): 514-5.
- Mølbak, K. and J. Neimann (2002). "Risk factors for sporadic infection with salmonella enteritidis, Denmark, 1997-1999." American Journal of Epidemiology **156**(7): 654-661.
- Moury, F., S. Fremy, et al. (2006). "Épidémiosurveillance des salmonelles d'origine non humaine : données du réseau salmonella - année 2005." Bulletin épidémiologique - AFSSA **20**: 6-8.
- Mullner, P., G. Jones, et al. (2009). "Source attribution of food-borne zoonoses in New Zealand: A modified hald model." Risk Analysis **29**(7): 970-984.
- Oliver, S. P., B. M. Jayarao, et al. (2005). "Foodborne pathogens in milk and the dairy farm environment: food safety and public health implications." Foodborne Pathog Dis **2**(2): 115-29.

- Peters, T. M., C. Maguire, et al. (2003). "The Salm-gene project - a European collaboration for DNA fingerprinting for." Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin **8**(2): 46-50.
- Pires, S. M., E. G. Evers, et al. (2009). "Attributing the human disease burden of foodborne infections to specific sources." Foodborne Pathog Dis **6**(4): 417-24.
- Pires, S. M., G. Nichols, et al. (2008). *Salmonella* source attribution in different European countries. 21st International ICFMH Symposium: "Evolving Microbial Food Quality and Safety". Aberdeen.
- Poirier, E., L. Watier, et al. (2008). "Evaluation of the impact on human salmonellosis of control measures targeted to *Salmonella* Enteritidis and Typhimurium in poultry breeding using time-series analysis and intervention models in France." Epidemiol Infect **136**(9): 1217-24.
- Popoff, M. Y. and L. Le Minor (1997). Antigenic formula of the *Salmonella* serovars, 7th revision. Paris, France, WHO Collaborating Center for Reference and Research on *Salmonella*. Institut Pasteur.
- Rabatsky-Ehr, T., J. Whichard, et al. (2004). "Multidrug-resistant strains of *Salmonella* enterica Typhimurium, United States, 1997-1998." Emerg Infect Dis **10**(5): 795-801.
- Rabsch, W., H. Tschäpe, et al. (2001). "Non-typhoidal salmonellosis: Emerging problems." Microbes and Infection **3**(3): 237-247.
- Rostagno, M. H., H. S. Hurd, et al. (2005). "Resting pigs on transport trailers as an intervention strategy to reduce *Salmonella* enterica prevalence at slaughter." Journal of Food Protection **68**(8): 1720-1723.
- Ruiz, M., J. C. Rodriguez, et al. (2004). "Available options in the management of non-typhi *Salmonella*." Expert Opinion on Pharmacotherapy **5**(8): 1737-1743.
- Sabuncu, E., J. David, et al. (2009). "Significant reduction of antibiotic use in the community after a nationwide campaign in France, 2002-2007." PLoS Medicine **6**(6).
- Sahlstrom, L., B. de Jong, et al. (2006). "*Salmonella* isolated in sewage sludge traced back to human cases of salmonellosis." Lett Appl Microbiol **43**(1): 46-52.
- Sarwari, A. R., L. S. Magder, et al. (2001). "Serotype distribution of *Salmonella* isolates from food animals after slaughter differs from that of isolates found in humans." Journal of Infectious Diseases **183**(8): 1295-1299.
- Singer, R. S., L. A. Cox Jr., et al. (2007). "Modeling the relationship between food animal health and human foodborne illness." Preventive Veterinary Medicine **79**(2-4): 186-203.
- Sirinavin, S. and P. Garner (2000). "Antibiotics for treating *Salmonella* gut infections." Cochrane Database Syst Rev. (2): CD001167.
- Sofos, J. N. (2008). "Challenges to meat safety in the 21st century." Meat Science **78**(1-2): 3-13.
- Souillard, R., J. Toux, et al. (2007). "Le RNOEA : Réseau National d'Observations Epidémiologiques en Aviculture Pathologie aviaire en 2004." Bulletin Epidémiologique - AFSSA **24**: 6-7.
- Stark, K. D., G. Regula, et al. (2006). "Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches." BMC Health Serv Res **6**: 20.
- Stoycheva, M. V. and M. A. Murdjeva (2006). "Antimicrobial therapy of salmonellosis--current state and perspectives." Folia medica. **48**(1): 5-10.
- Swaminathan, B., T. J. Barrett, et al. (2001). "PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States." Emerging infectious diseases **7**(3): 382-389.
- Swanson, S. J., C. Snider, et al. (2007). "Multidrug-resistant *Salmonella* enterica serotype Typhimurium associated with pet rodents." N Engl J Med **356**(1): 21-8.
- Tauxe, R. V. (1997). "Emerging Foodborne Diseases: An Evolving Public Health Challenge." Emerging Infectious Diseases **3**(4): 425-434.

- Tindall, B. J., P. A. D. Grimont, et al. (2005). "Nomenclature and taxonomy of the genus *Salmonella*." International Journal of Systematic and Evolutionary Microbiology **55**(1): 521-524.
- Todd, E. C., J. D. Greig, et al. (2008). "Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 5. Sources of contamination and pathogen excretion from infected persons." J Food Prot **71**(12): 2582-95.
- Tsen, H.-Y. (2002b). "Molecular typing of *Salmonella enterica* serovars Typhimurium, Typhi, and Enteritidis isolated in Taiwan." Journal of Food and Drug Analysis **10**(4): 242-251.
- Uyttendaele, M. R., J. M. Debevere, et al. (1998). "Prevalence of *Salmonella* in poultry carcasses and their products in Belgium." International Journal of Food Microbiology **40**(1-2): 1-8.
- Vaillant, V. and E. Espie (2005). "Cases of *Salmonella* Stourbridge infection in France, April-June 2005." Euro Surveill **10**(7): E050721 2.
- Varma, J. K., K. Mølbak, et al. (2005). "Antimicrobial-resistant nontyphoidal *Salmonella* is associated with excess bloodstream infections and hospitalizations." Journal of Infectious Diseases **191**(4): 554-561.
- Velge, P., A. Cloeckaert, et al. (2005). "Emergence of *Salmonella* epidemics: The problems related to *Salmonella enterica* serotype Enteritidis and multiple antibiotic resistance in other major serotypes." Veterinary Research **36**(3): 267-288.
- Voetsch, A. C., T. J. Van Gilder, et al. (2004). "FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States." Clin Infect Dis **38 Suppl 3**: S127-34.
- Volatier, J. L. (2000). Enquête nationale sur les consommations alimentaires. Tec&Doc, CREDOC-AFSSA-DGAL.
- Wegener, H. C., T. Hald, et al. (2003). "*Salmonella* control programs in Denmark." Emerging Infectious Diseases **9**(7): 774-780.
- Weill, F.-X., F. Guesnier, et al. (2006). "Multidrug resistance in *Salmonella enterica* serotype typhimurium from humans in France (1993 to 2003)." Journal of Clinical Microbiology **44**(3): 700-708.
- Weill, F. X. (2008). "*Salmonella* : épidémiologie, typage et résistance aux antibiotiques." Revue Francophone des Laboratoires **400**: 37-47.
- Weinberger, M., N. Andorn, et al. (2004). "Blood invasiveness of *Salmonella enterica* as a function of age and serotype." Epidemiol Infect **132**(6): 1023-8.
- Wilson, D. J., E. Gabriel, et al. (2008). "Tracing the source of campylobacteriosis." PLoS Genet **4**(9): e1000203.
- Woo, Y.-K. and S.-H. Lee (2006). "Genetic diversity of multi-resistant *Salmonella enterica* serotype typhimurium isolates from animals and humans." Journal of Microbiology **44**(1): 106-112.
- Woodward, D. L., R. Khakhria, et al. (1997). "Human salmonellosis associated with exotic pets." Journal of Clinical Microbiology **35**(11): 2786-2790.
- Zheng, J., C. E. Keys, et al. (2007). "Enhanced subtyping scheme for *Salmonella enteritidis*." Emerging Infectious Diseases **13**(12): 1932-1935.
- Zwietering, M. H. (2008). "Quantitative risk assessment: Is more complex always better? Simple is not stupid and complex is not always more correct." Int J Food Microbiol.

Annexes

Annexe 1

L'inférence bayésienne

1 Pourquoi le raisonnement bayésien

Rien n'est jamais vraiment objectif. Dans le mode de raisonnement scientifique « classique », la formulation d'une hypothèse se fait selon les convictions et connaissances antérieures du scientifique, de même pour le choix des conditions initiales ou limites, de la limite de significativité ou encore des points extrêmes à exclure. De plus, les opinions des scientifiques influencent leurs conclusions. Ainsi face à un même phénomène objectif, différents scientifiques tireront des conclusions différentes à partir de données pourtant identiques.

La subjectivité n'est donc pas un obstacle ou une erreur mais une composante inhérente aux statistiques et à la méthodologie scientifique. Elle est même essentielle et incontournable pour la création de nouvelles connaissances. Cependant, la traiter à la légère (lui donner trop de place ou ne pas la traiter) peut être source d'erreur, d'où la nécessité d'une approche scientifique formalisée de la subjectivité. La méthode bayésienne permet ainsi d'intégrer formellement la subjectivité jusqu'ici sous-jacente aux raisonnements scientifiques. Elle intègre de plus scientifiquement le phénomène d'apprentissage, qui se caractérise classiquement par rejet d'hypothèse, formulation d'une nouvelle hypothèse et modification de l'expérimentation pour tester la nouvelle hypothèse.

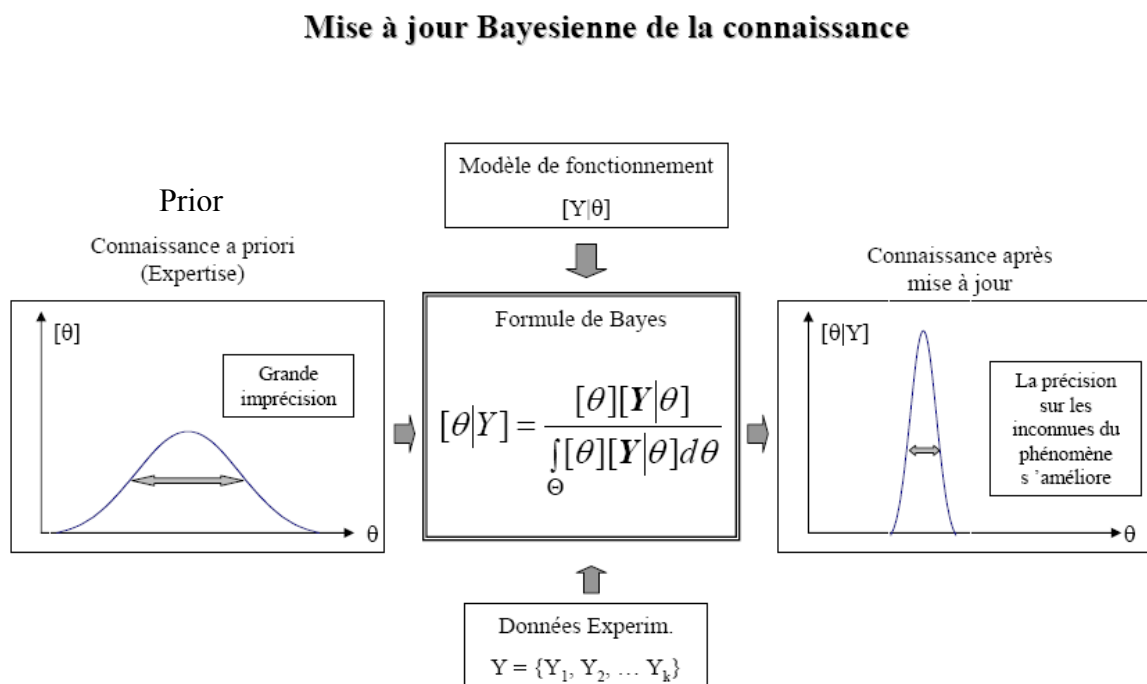
Enfin, quant à l'intérêt d'intégrer la part de subjectivité que constituent la culture scientifique et les connaissances acquises : Bernouilli dans ARS coniectandi

« Pour une composition donnée de l'air et pour des masses, positions, directions et vitesse des vents donnés, pour une composition en vapeur et des nuages donnés, et connaissant les lois de la mécanique qui gouvernent toutes ces interactions, le temps de demain ne sera pas différent de ce qu'il devrait en fait être. [...] Et si nous considérons cet événement comme un événement aléatoire, c'est que nos connaissances des lois gouvernant la succession des événements climatiques, mais aussi de la mécanique et des mathématiques, sont insuffisantes pour calculer et prédire le phénomène (alors que nous sommes capables de prédire les éclipses connaissant les principes absolus de l'astronomie). Ce qui relève du hasard dépend donc finalement du niveau de nos connaissances. »

2 Principe

Mise à jour bayésienne de la connaissance : on apprend en modifiant ce que l'on savait déjà, à partir de ce que l'on observe.

L'inférence bayésienne est un calcul a posteriori combinant savoir a priori et information apportée par les données. Elle est portée par le principe de pari sur les valeurs inconnues du paramètre, le conditionnement se faisant par rapport aux données ou apprentissage statistique. L'estimation se fait par répartition de probabilités.



In Modélisation markovienne pour l'évaluation et la gestion des risques, Eric Parent, 2004

Figure 1 : Principe de la modélisation bayésienne

2.1 Prior : partie subjective

Le prior correspond au degré de croyance qu'a l'analyste avant d'observer la moindre donnée en relation avec le problème.

Il est à noter que dans les cas où aucune donnée n'est disponible (manque de temps, expérimentation trop coûteuse, ...), le prior sera utilisé seul pour l'inférence et la prise de décisions. D'où l'importance d'une formulation soignée. Enfin, l'influence du prior sur la

probabilité postérieure est relativisée par la taille de l'échantillon de données. Plus le nombre de données sera grand, moins le prior pèsera sur la probabilité postérieure et inversement.

2.2 Observations : partie objective

Elles sont intégrées au modèle par une fonction de vraisemblance qui résume l'information objective. Le terme de vraisemblance semble avoir été utilisé pour la première fois par Fisher en 1921 ; il entendait par là une conviction rationnelle sur une quantité inconnue. La fonction de vraisemblance n'est pas une probabilité et n'obéit pas aux lois de la probabilité. Elle correspond à la fonction de densité jointe d'un ensemble d'observations, vue comme une fonction du paramètre inconnu θ .

$$L(\theta; y_1, \dots, y_n) = f(y_1, \dots, y_n / \theta) \quad \text{pour tout } \theta.$$

Elle est définie de façon unique à une constante multiplicative près, constante qui ne joue aucun rôle dans le calcul du maximum de vraisemblance. Fisher utilisait cette notion comme mesure de la plausibilité d'un θ , grâce au ratio de plausibilité : $L(\hat{\theta})/L(\theta)$, où $\hat{\theta}$ correspond au maximum de vraisemblance.

Dans le raisonnement bayésien, la fonction de densité postérieure ne dépend des données qu'au travers de la fonction de vraisemblance.

$$h(\theta / y_1, \dots, y_n) = \frac{cL(\theta; y_1, \dots, y_n)g(\theta)}{\int cL(\theta; y_1, \dots, y_n)g(\theta)d\theta'}$$

Ainsi, définir correctement la vraisemblance est fondamental en raisonnement bayésien, mais celle-ci ne pourra être approchée précisément que dans le cadre d'échantillons importants, elle sera source (comme le prior) d'incertitudes dans le cas de petits échantillons ou d'échantillons moyens.

2.3 Inférence statistique

Celle-ci repose sur le théorème de Bayes qui est une simple assertion de probabilités conditionnelles.

Soit A_1, \dots, A_k un ensemble d'événements mutuellement exclusifs et exhaustifs, dont A_j est d'un intérêt particulier, ainsi que l'événement B ,

$$P\{A_j / B\} = \frac{P\{B / A_j\}P\{A_j\}}{\sum_{i=1}^k P\{B / A_i\}P\{A_i\}}, \text{ pour } P(B) \neq 0.$$

$P\{A_j\}$: prior, probabilité personnelle de l'événement A_j , degré de croyance.

$P\{A_j / B\}$: probabilité postérieure, degré de croyance après avoir les résultats de l'événement B.

Autrement dit, ce théorème fournit un moyen de mettre à jour le degré de croyance d'un événement A_j au vu des nouvelles informations B portant sur A_j . C'est une loi intégrant la notion d'apprentissage par l'expérience.

2.4 Probabilité postérieure

Elle résulte de l'application du théorème. La probabilité postérieure pourra servir de prior pour une expérimentation ultérieure, dans la logique d'apprentissage qui sous-tend le bayésien. Elle exprime l'incertitude sur l'inconnue θ après observations des données.

Annexe 2

Synthèse du protocole des études communautaires de référence Règlement européen (EC) 2160/2003

1 Plan d'échantillonnage et prélèvements

1.1 Poules pondeuses

L'étude poules pondeuses concernait les élevages de poules pondeuses d'œufs de consommation comprenant au moins 1 000 poules pondeuses commercialisant tout ou partie de leur production via un centre d'emballage. Elle a été menée du 1^{er} octobre 2004 au 30 septembre 2005 et a porté sur 518 exploitations désignées par un tirage au sort stratifié sur la taille d'élevage dans la base de données 2003 des exploitations françaises de poules pondeuses constituée sur déclaration des DDSV (Note de service N2004-8227). Cette taille minimale d'échantillon a été déterminée pour l'estimation au risque de 5% d'une prévalence *a priori* de 20% avec une précision absolue de 3%. Les exploitations initialement tirées au sort ont pu être remplacées en cas de non-conformité aux critères d'inclusion. Un troupeau est étudié par exploitation au cours des 9 semaines qui précèdent la fin de vie économique prévue de ce troupeau et après 60 semaines d'âge. Un questionnaire est rempli avec l'éleveur et 7 prélèvements sont réalisés par les agents des services vétérinaires, leur nature dépend du mode d'élevage (Cf table 1).

Mode d'élevage	Prélèvements	
	Matières fécales	Poussières
Cage	5 prélèvements de fientes prélevés du tapis de collecte des fientes, du racloir ou de la fosse, suivant le type d'aménagement du bâtiment	2 échantillons de poussières prélevés au pied des cages ou sous les cages.
Sol Plein air	5 prélèvements de fientes prélevés à l'aide de 5 paires de pédichiffonnettes ou de stérilottes.	1 échantillon de poussières prélevées du collecteur d'œufs. 1 échantillon de poussières prélevées en différents points du bâtiment.

Table 1 : Nature des prélèvements selon le mode d'élevage pour l'étude communautaire de référence concernant les poules pondeuses (Note de service N2004-8241)

1.2 Poulet de chair

L'étude poulets de chair concernait les exploitations d'une capacité supérieure à 5 000 poulets de chair. Elle a été menée du 1^{er} octobre 2005 au 30 septembre 2006 et a porté sur 371 exploitations désignées par un tirage au sort stratifié sur la taille d'élevage dans la base de données 2003 des exploitations françaises de poulets de chair constituée sur déclaration des DDSV. Comme pour l'étude dinde (figure 1), les exploitations initialement tirées au sort ont pu être remplacées en cas de non-conformité aux critères d'inclusion. Un troupeau est étudié par exploitation au cours de la semaine précédent le dernier abattage du troupeau. Un questionnaire est rempli avec l'éleveur et 5 prélèvements sont réalisés par les agents des services vétérinaires (Cf table 2).

Mode d'élevage	Prélèvements
	Matières fécales
Sol Plein air	5 prélèvements de fientes, recueillis à l'aide de 5 paires de stérilottes humidifiées à l'intérieur du bâtiment d'élevage.

Table 2 : Nature des prélèvements pour l'étude communautaire de référence concernant les poulets de chair (Note de service N2005-8240)

1.3 Dinde

L'étude dinde concernait les exploitations de production d'une capacité supérieure ou égale à 500 dindes de chair présentes. Elle a été menée du 1^{er} octobre 2005 au 30 septembre 2006 et a porté sur 374 exploitations désignées par un tirage au sort stratifié sur la taille d'élevage dans la base de données SIGAL. Cette taille minimale d'échantillon a été déterminée pour l'estimation au risque de 5% d'une prévalence *a priori* de 50% avec une précision absolue de 5% (figure 1). Les exploitations initialement tirées au sort ont pu être remplacées en cas de non-conformité aux critères d'inclusion. Un troupeau est étudié par exploitation dans la semaine précédant son abattage. Un questionnaire est rempli avec l'éleveur et 5 prélèvements sont réalisés par les agents des services (Cf table 3).

Mode d'élevage	Prélèvements = matières fécales
sol ou plein-air	<u>pour chaque troupeau</u> : prélèvements de matières fécales recueillies à l'aide de 5 paires de sterisocks humidifiées, à l'intérieur du bâtiment d'élevage.

Table 3 : Nature des prélèvements selon le mode d'élevage pour l'étude communautaire de référence concernant les dindes (Note de service N2006-8255)

1.4 Porcs

L'étude porc concernait les porcs charcutiers de 50 à 170 kg engraisés et abattus en France dans les 21 établissements abattant 80% de la production nationale. Elle a été menée du 1^{er} octobre 2006 au 30 septembre 2007 et a porté sur 1 176 carcasses désignées par un tirage au sort simple stratifié par mois et par abattoir participant, proportionnel aux capacités d'abattage des abattoirs. Cette taille minimale d'échantillon a été déterminée pour l'estimation au risque de 5% d'une prévalence *a priori* de 50% avec une précision absolue de 5%, en incluant un échantillon supplémentaire de 10%. Pour un jour ouvré, un numéro de lot et un numéro d'ordre de la carcasse au sein du lot sont déterminés par tirage aléatoire simple. Les prélèvements à réaliser sur les carcasses sont les suivants (Cf table 4).

Type d'échantillon	Nombre de carcasses à prélever	Détail des prélèvements
Ganglions lymphatiques mésentériques	1176	<i>Pour chaque carcasse sélectionnée</i> : au moins 4 ganglions lymphatiques iléaux distaux et tous les ganglions iléo-cæcaux seront prélevés, afin que l'échantillon collecté représente une masse d'au moins 25 g de tissu lymphatique <u>sans gras ni tissus conjonctifs</u> . (compléter si nécessaire en prélevant une partie de la chaîne lymphatique jéjunale distale).
Echantillon musculaire destiné à l'obtention de jus de viande	1176	<i>Pour chaque carcasse sélectionnée</i> : prélever assez de muscles (environ 15-20 g selon la consistance) sur le cou ou le diaphragme afin d'en obtenir une quantité de jus de viande suffisante pour en congeler et en stocker une partie.
Prélèvements de surface de carcasses	424	<i>Pour chaque carcasse sélectionnée</i> : après l'éviscération mais avant le refroidissement, une seule éponge est frottée sur les 4 sites suivants : <ul style="list-style-type: none">- membre postérieur, partie médiale ;- abdomen, partie latérale ;- région dorsale médiane ;- joue. Les mêmes carcasses font systématiquement l'objet d'un prélèvement d'amygdales.

Table 4 : Nature des prélèvements concernant la recherche de salmonelles pour l'étude communautaire de référence concernant les porcs (Note de service N2006-8293)

Les prélèvements sont envoyés pour analyse bactériologiques au LNR-Salmonella de Ploufragan. Les échantillons musculaires sont destinés à une recherche sérologique, que nous n'utiliserons pas ici.

2 Traitement des échantillons

Pour les poules pondeuses, poulets de chair et dindes, les prélèvements sont envoyés pour analyse bactériologiques au LNR-Salmonella de Ploufragan. Une souche est isolée par prélèvement après enrichissement (protocole dérivé de la norme ISO 6579 (2002)) et sérotypée. Un élevage est considéré comme positif si l'un au moins des échantillons est positif.

Pour les porcs, une souche est isolée par prélèvement de type ganglions lymphatiques mésentériques ou surface de carcasses après enrichissement et analysée de la même façon

que précédemment. Le traitement des données se fait de façon indépendante pour les deux types d'échantillons, une carcasse est considérée positive si une salmonelle est isolée dans l'échantillon.

3 Exemple de plan d'échantillonnage : étude communautaire de référence dindes

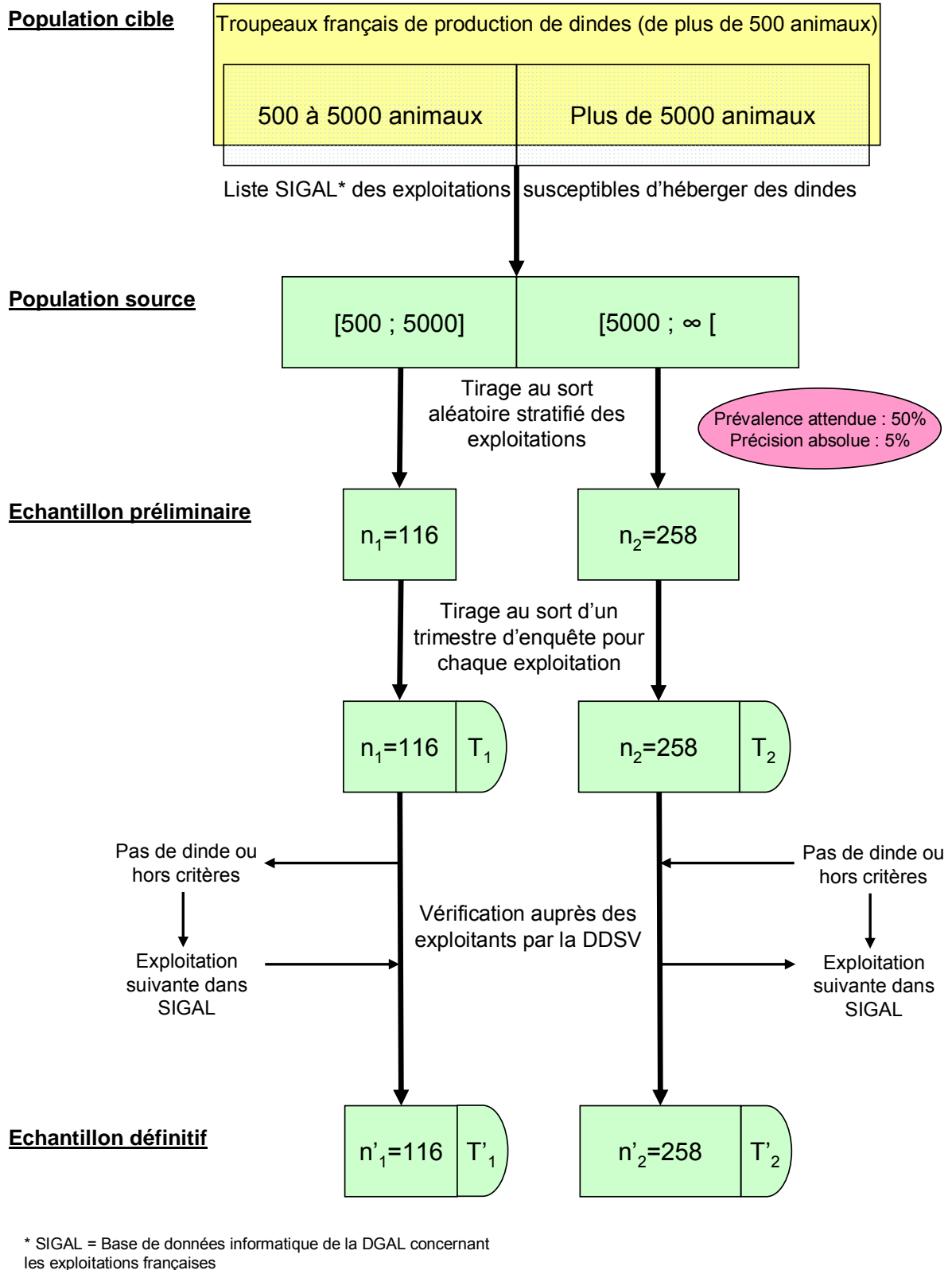


Figure 1 : Plan d'échantillonnage pour l'étude communautaire de référence en élevage de production de dindes

Annexe 3

Protocole du plan national de surveillance de l'antibiorésistance de certaines bactéries sentinelles et zoonotiques chez les bovins

1 Plan d'échantillonnage et prélèvements

Ce plan porte sur les carcasses issues de 3 filières : veaux de boucherie, jeunes bovins et réforme d'allaitantes. Il a été mené pour 2005 de septembre à décembre dans 3 départements parmi les plus gros producteurs pour chaque filière : Ille-et-Vilaine, Dordogne et Pyrénées-Atlantiques pour les veaux, Maine-et-Loire, Saône-et-Loire et Moselle pour les jeunes bovins et Allier, Tarn et Haute-Vienne pour les réformes allaitantes. Ce plan a porté sur 334 carcasses réparties également entre les filières. Les prélèvements sont effectués en abattoir en fonction de la quantité produite sur chaque site et la répartition entre les départements est réalisée en vue d'obtenir un échantillonnage globalement représentatif de la production française. Un animal est testé au plus par exploitation et les lots à prélever sont choisis de façon aléatoire. 35g de fécès sont prélevés dans le rectum par les agents de la DDSV.

2 Traitement des échantillons

Les prélèvements sont envoyés pour analyse bactériologiques au laboratoire de l'AFSSA de Lyon. Les salmonelles sont recherchées sur enrichissement, puis sérotypées et antibiotypées par CMI (Concentration maximale inhibitrice).

Annexe 4

Données de surveillance animale active

Prévalences et répartitions par sérotype selon les filières

	Prévalence	Répartition
spp	30,82%	100,00%
Typhimurium	6,89%	19,13%
Infantis	5,49%	6,96%
Anatum	5,25%	6,09%
Enteritidis	4,09%	17,39%
Braenderup	3,13%	5,22%
Bareilly	2,95%	0,87%
Mbandaka	2,81%	6,96%
Tennessee	2,37%	5,22%
Livingstone	1,98%	4,35%
Montevideo	1,92%	2,61%
Agona	0,98%	3,48%
Virchow	0,86%	2,61%
Coeln	0,73%	0,87%
Llandoff	0,72%	1,74%
Yobura	0,36%	0,87%
S I 1,3,19:z27:-	0,29%	0,87%
Schwarzengrund	0,29%	0,87%
Senftenberg	0,27%	0,87%
Jerusalem	0,15%	2,61%
Reading	0,15%	0,87%
Lexington	0,15%	0,87%
Gloucester	0,11%	0,87%
Oranienburg	0,09%	0,87%
Hadar	0,08%	0,87%
Holcomb	0,07%	0,87%
Szentes	0,07%	0,87%
S I 4,5:i:-	0,06%	0,87%
Havana	0,04%	0,87%
S I 16:d:-	0,04%	0,87%
S I 6,7:d:-	0,03%	0,87%
SII 6,7:z:1,5	0,02%	0,87%

Table 1 : Prévalence et répartition globales et par sérotype pour les poules pondeuses (données DGAL, 2004-2005)

	Prévalence	Répartition
spp	8,23%	100%
Hadar	1,55%	8,82%
Virchow	1,53%	8,82%
Montevideo	1,31%	8,82%
Mbandaka	0,88%	5,88%
Newport	0,49%	2,94%
Schwarzengrund	0,49%	2,94%
Kedougou	0,45%	2,94%
Enteritidis	0,37%	5,88%
Indiana	0,34%	5,88%
Paratyphib	0,33%	2,94%
Anatum	0,21%	8,82%
Braenderup	0,14%	2,94%
Livingstone	0,07%	2,94%
Infantis	0,07%	2,94%
Senftenberg	0,07%	2,94%
Lexington	0,07%	2,94%
Typhimurium	0,07%	2,94%
Ohio	0,07%	2,94%
lii b k z	0,07%	2,94%
Napoli	0,07%	2,94%
Lii	0,07%	2,94%
Heidelberg	0,02%	2,94%
Veneziana	0,02%	2,94%

Table 2 : Prévalence et répartition globales et par sérotype pour les poulets de chair (données DGAL, 2005-2006)

	Prévalence	Répartition
spp	13,78%	100,00%
Derby	3,55%	25,00%
Typhimurium	2,63%	15,38%
Mbandaka	1,58%	9,62%
Agona	0,82%	5,77%
Newport	0,77%	3,85%
Hadar	0,76%	5,77%
Indiana	0,75%	5,77%
Bredeney	0,72%	3,85%
Enteritidis	0,71%	11,54%
Saintpaul	0,63%	3,85%
Napoli	0,42%	1,92%
Muenster	0,38%	1,92%
Remo	0,17%	1,92%
Coeln	0,11%	1,92%
Aijobo	0,04%	1,92%

Table 3 : Prévalence et répartition globales et par sérotype pour les dindes (données DGAL, 2006-2007)

	Prévalence	Répartition
spp	18,18%	100,00%
Typhimurium	7,55%	41,51%
Derby	6,52%	35,85%
Agona	0,43%	2,36%
Infantis	0,43%	2,36%
Brandenburg	0,34%	1,89%
Schwarzengrund	0,34%	1,89%
Bredeney	0,26%	1,42%
S I 4,12:-:1,2	0,26%	1,42%
Anatum	0,17%	0,94%
Enteritidis	0,17%	0,94%
Kedougou	0,17%	0,94%
Newport	0,17%	0,94%
Bovismorbificans	0,09%	0,47%
Bradford	0,09%	0,47%
Coeln	0,09%	0,47%
Goldcoast	0,09%	0,47%
Hadar	0,09%	0,47%
Indiana	0,09%	0,47%
Lercap	0,09%	0,47%
Mbandaka	0,09%	0,47%
Montevideo	0,09%	0,47%
S III 48:z4,z23:-	0,09%	0,47%
S IIIa 48:z4,z23:-	0,09%	0,47%
S IV 40:z4,z23:-	0,09%	0,47%
Saint Paul	0,09%	0,47%
Stourbridge	0,09%	0,47%
Tennessee	0,09%	0,47%
Veneziana	0,09%	0,47%

Table 4 : Prévalence et répartition globales et par sérotype pour les porcs (données DGAL, 2006-2007)

	Prévalence	Répartition
spp	2,4%	100,0%
Mbandaka	1,5%	62,5%
rough	0,3%	12,5%
Salamae	0,3%	12,5%
Indiana	0,3%	12,5%

Table 5 : Prévalence et répartition globales et par sérotype pour les porcs (données DGAL, 2006-2007)

Annexe 5

Données de surveillance animale passive Répartitions par sérotype selon les filières

Sérotype	Pondeuse	Dinde	Poulet de chair	Bovin	Porc	Ovin	Produits de la mer	Canard	Autres volailles
Enteritidis	27,2%	2,8%	17,8%	2,8%	0	3,8%	0	1,5%	2,2%
Senftenberg	19,8%	14,7%	13,0%	0	0	0	0	1,3%	2,8%
Typhimurium	12,6%	15,5%	9,7%	22,5%	23,5%	6,7%	15,0%	23,6%	38,5%
Montevideo	3,7%	1,0%	2,4%	14,1%	0,5%	0	7,5%	7,0%	2,9%
Mbandaka	3,1%	0,5%	0,9%	1,4%	0,9%	0	0	0,1%	1,0%
Indiana	2,3%	9,5%	10,4%	0	0,5%	0	2,5%	23,1%	31,0%
Infantis	1,5%	1,0%	3,9%	2,8%	6,6%	0	5,0%	4,9%	0,6%
Virchow	1,5%	0,3%	5,9%	0	0	0	2,5%	0,1%	0,1%
Anatum	1,0%	0,5%	1,3%	2,8%	0	0	5,0%	0,8%	0,6%
Hadar	0,8%	8,2%	5,5%	0	1,4%	0	0	1,7%	2,5%
Agona	0,8%	8,3%	2,6%	1,4%	2,3%	0	0	0,3%	1,1%
Derby	0,6%	9,0%	2,2%	19,7%	16,4%	3,8%	5,0%	0,1%	1,9%
Coeln	0,5%	2,4%	0,2%	1,4%	0	0	0	0	0,4%
Kottbus	0,3%	3,8%	2,7%	0	0,5%	0	0	23,2%	1,8%
Bredeney	0,3%	6,9%	1,0%	0	0,9%	0	0	0,1%	2,5%
Napoli	0,2%	3,8%	0,2%	0	0	0	0	0,7%	0,6%
Brandenburg	0,2%	0,9%	0,2%	4,2%	2,8%	0	0	0	0
Dublin	0,2%	0	0	2,8%	0	0	0	0	0
Kentucky	0,2%	0	0,5%	7,0%	0	0	0	0	0,3%
Oranienburg	0,2%	0	0	0	0	0	0	0	0
Heidelberg	0,2%	0,6%	3,0%	0	0	0	0	0	0,3%
Muenster	0,2%	0,8%	0,1%	0	0	0	2,5%	0	0
Newport	0	1,6%	0,6%	0	0	1,9%	0	0,3%	2,8%
Manhattan	0	0	0	0	7,5%	0	0	0	0
Stourbridge	0	0	0	0	0	0	0	0	0,1%
Bovismorbificans	0	0,5%	0	0	0,9%	0	0	0	0
Blockley	0	0	0,1%	0	0	0	0	0	0
Poona	0	0	0	0	0	0	2,5%	0	0
Weltevreden	0	0	0	0	0	0	25,0%	0	0
Sandiego	0	0	0	0	0	0	0	0,1%	0
S 61:k:1,5,7	0	0	0	0	0	57,7%	0	0	0
Autres	22,9%	7,5%	15,8%	16,9%	35,2%	26,0%	27,5%	11,0%	6,1%
Effectif total	371	651	924	53	163	93	34	2330	428

Table 1 : Répartition des souches par sérotype et par filière en surveillance passive

Annexe 6

Méthode de classification des souches

L'ACM (analyse en correspondances multiples) est une analyse multidimensionnelle (Lebart, Morineau et al. 2000) permettant de décrire des données nominales (qualitatives ou quantitatives en classes) dont on ne connaît pas la structure. Elle permet de représenter les individus dans un espace facile à appréhender, c'est-à-dire en 2 ou 3 dimensions, en minimisant la perte d'information. Ceci se fait en déterminant un petit nombre de variables (quantitatives) qui sont une combinaison linéaire des variables initiales et qui ont la propriété de traduire la plus grande dispersion des individus et donc l'hétérogénéité des données. Sur un graphique d'ACM, les individus les plus éloignés du centre sont les plus originaux et la distance entre deux individus sera d'autant plus importante qu'ils ont des modalités différentes pour de nombreuses variables et que ces modalités sont rares. Chacun des individus (ici les souches) est caractérisé par des coordonnées quantitatives continues dans le nouvel espace constitué des axes (ou variables) défini(e)s à partir des données initiales.

On utilise alors une classification de type mixte, qui correspond à l'utilisation conjointe de la méthode d'agrégation autour de centres mobiles et de la classification ascendante hiérarchique, en utilisant le critère de Ward. Dans une première étape, les souches subissent un partitionnement par l'algorithme des centres mobiles (constitution de nombreuses classes homogènes) puis les classes formées sont agrégées selon la méthode de classification ascendante hiérarchique, sur le critère de Ward (minimisation de la variabilité intra-classe et maximisation de la variabilité inter-classes). Cette méthode conduit à constituer un dendrogramme de partitions possibles à partir de k groupes initiaux. Enfin, les partitions correspondant aux coupures initialement choisies sont optimisées, à nouveau par l'algorithme des centres mobiles. Cette dernière technique est la suivante : on choisit d'abord un nombre n de classes et on choisit aléatoirement n individus qui seront les noyaux initiaux de classes. Puis, chaque individu est affecté au noyau le plus proche (en distance euclidienne). On obtient ainsi n classes provisoires. Les n noyaux initiaux sont alors remplacés par le barycentre des n classes provisoires et on recommence l'opération selon un processus itératif d'affectation des individus à la classe la plus proche jusqu'à ce que les classes soient stables ou que la variabilité intra-classe cesse de décroître de façon sensible. Ceci nous permet d'obtenir plusieurs propositions de classification avec un nombre de clusters variable. Les critères que nous avons utilisés pour choisir la partition optimale ont été proposés par Berge (Berge, Atwill et al. 2003). La variabilité décrite par les clusters augmente avec le nombre de clusters, il s'agit donc ici d'optimiser le niveau de variabilité

décrit sans créer trop de clusters. Si au sein d'un cluster au moins, la réponse est non uniforme à plus de deux classes d'antibiotiques, on choisit une partition avec un nombre de clusters plus important. En revanche, si les effectifs dans au moins un cluster sont inférieurs à 2% de l'effectif total, on choisit une partition avec un plus petit nombre de clusters. Enfin, concernant la signification des clusters, nous avons essentiellement considéré la présence simultanée de souches humaines et animales, afin d'éviter de dissocier souches humaines et animales, les autres critères étant satisfaits par ailleurs. Cette méthode doit permettre d'obtenir une classification minimisant le nombre de clusters à faible effectif, parcimonieuse en nombre de clusters et minimisant la variabilité intra-cluster, et ainsi réduire les biais potentiels de classification.

1. Lebart, L., A. Morineau, and M. Piron, *Statistique exploratoire multidimensionnelle*. 3ième édition ed. 2000, Paris: Dunod. 439.
2. Berge, A.C.B., E.R. Atwill, and W.M. Sisco, *Assessing antibiotic resistance in fecal Escherichia coli in young calves using cluster analysis techniques*. Preventive Veterinary Medicine, 2003. **61**(2): p. 91-102.

Attribution des cas de salmonelloses humaines aux différentes filières de production animale en France

Adaptabilité et robustesse du modèle bayésien d'attribution par typage microbiologique

Résumé

Les salmonelles sont la première cause bactérienne d'infections d'origine alimentaire en France. Ces pathogènes zoonotiques sont présents dans de nombreuses sources animales et font l'objet d'une lutte dès l'élevage. Ainsi, identifier la contribution relative des différentes sources aux cas humains est essentiel pour mettre en place une politique de lutte efficace. L'outil bayésien d'attribution par typage microbiologique qui permet de déterminer le nombre de cas liés à chaque source est pour cela particulièrement adapté. Le principe du modèle consiste à comparer la distribution de types microbiologiques (sérotypes et sous-types), parmi les cas humains et dans les sources animales, en pondérant par les quantités des sources consommées et en tenant compte des spécificités des sources et des types. Les objectifs des travaux de thèse étaient d'identifier et collecter les données nécessaires pour appliquer un tel modèle et d'évaluer sa robustesse et l'impact de la qualité des données sur les résultats. Le système de surveillance français des salmonelles a pour cela été décrit et analysé, et une base de données suffisante pour appliquer le modèle a été constituée. L'application du modèle dans divers pays européens a montré les limites de l'approche. Comme attendu, le modèle est très sensible à l'information *a priori* informative que sa structure surparamétrée rend nécessaire. Nous avons proposé une paramétrisation plus adaptée basée sur les données. L'application du modèle à des données de surveillance passive (proportions) et actives (prévalences) a alors montré la robustesse du modèle. Ainsi, mettre en place une démarche d'attribution est possible en France. Les limites méthodologiques sont à explorer plus avant, mais les perspectives d'utilisation, notamment pour attribuer les cas en fonction de la résistance aux antibiotiques, justifient les efforts requis.

Mots-clés : Attribution, *Salmonella*, santé publique, modélisation bayésienne, risque alimentaire, épidémiologie, zoonoses

Abstract

Salmonella are the first bacterial cause of food-borne infections in France. These zoonotic pathogens occur in a variety of animal sources and are the aim of interventions from the farm level on. Identification of the relative contribution of the different sources to human cases is essential to implement an efficient policy to fight these infections. The bayesian microbial subtyping attribution approach allows determining the number of cases related to a particular source, and is thus particularly appropriate to this aim. The principle of this model consists in comparing the distribution of microbiological types (serotypes or subtypes) in human cases and in animal sources, weighted by the amount of sources consumed and taking into account the specificity of both the sources and the serotypes. The objectives of the thesis work were to identify and collect the data that are necessary to run such a model and to assess its robustness and the impact of the quality of the data on the results. The French surveillance system for *Salmonella* was thus described and analysed and a database sufficient to allow using the model was built. The application of the model in different European Member States showed the limits of the approach. As expected, the model proves to be very sensitive to the prior informative information that its overparameterized structure requires. We proposed a more appropriate parameterization based on the data. The use of the model on surveillance data of both passive (proportions) and active (prevalence) nature then demonstrated the robustness of this model. Thus, the source attribution approach could be implemented in France. The methodological limits are to be further explored but the perspectives of use, particularly to attribute the cases according to antibiotic resistance, justify mustering the necessary means.

Keywords: source attribution, *Salmonella*, public health, Bayesian modeling, foodborne risk, epidemiology, zoonoses