Strategy for systematic review of observational studies and meta-analysis modelling of risk factors for sporadic foodborne diseases

Ursula Gonzales-Barron, Anne Thébault, Pauline Kooh, Laurence Watier, Moez Sanaa, Vasco Cadavez

▶ **To cite this version:**

Ursula Gonzales-Barron, Anne Thébault, Pauline Kooh, Laurence Watier, Moez Sanaa, et al.. Strategy for systematic review of observational studies and meta-analysis modelling of risk factors for sporadic foodborne diseases. Microbial Risk Analysis, 2021, 17, pp.100082. 10.1016/j.mran.2019.07.003 . anses-03436698

HAL Id: anses-03436698

https://anses.hal.science/anses-03436698

Submitted on 26 Nov 2021

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.

Microbial Risk Analysis xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Microbial Risk Analysis

journal homepage: www.elsevier.com/locate/mran

Full length article

Strategy for systematic review of observational studies and meta-analysis modelling of risk factors for sporadic foodborne diseases

Ursula Gonzales-Barron[a,*], Anne Thébault[b], Pauline Kooh[b], Laurence Watier[c], Moez Sanaa[b], Vasco Cadavez[a]

[a] CIMO Mountain Research Centre, School of Agriculture, Polytechnic Institute of Bragança, Portugal
[b] Department of Risk Assessment, French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Maisons-Alfort, France
[c] Department of Biostatistics, Biomathematics, Pharmacoepidemiology and Infectious Diseases (B2PHI), Institut National de la Santé et de la Recherche Médicale (INSERIM), UVSQ, Institut Pasteur, Université Paris-Saclay, Paris, France

ARTICLE INFO

Keywords:
Gastrointestinal disease
Case-control study
Cohort study
Exposure
Transmission
Odds-ratio

ABSTRACT

In order to design effective public health strategies, and, in particular, effective food safety interventions to reduce the burden of foodborne disease, the most important sources of enteric illnesses should be identified. Both case-control and cohort observational studies have for long been powerful approaches among epidemiologists to investigate the association of exposure and illness. In the literature, there are numerous case-control and cohort studies reporting results on risk factors and routes of transmission of sporadic foodborne infections. The objective of this article is to describe, in depth, the strategies implemented for systematic review and meta-analysis of the associations between multiple risk factors and eleven food and waterborne diseases, namely, non-typhoidal salmonellosis, campylobacteriosis, Shiga-toxin *E. coli* infection, listeriosis, yersiniosis, toxoplasmosis, norovirus infection, hepatitis A, hepatitis E, cryptosporidiosis and giardiasis. First, this article describes the procedures of systematic searches in five bibliographic engines, screening of relevance and assessment of methodological quality according to pre-set criteria. It proceeds with the explanation of a broad data categorisation scheme established to hierarchically group the risk factors into travel, host-specific factors and pathways of exposure (i.e., person-to-person, animal, environment and food routes), with views to harmonising and supporting the integration of outcomes from studies investigating a variety of potential determinants of disease. Next, the article describes the four meta-analysis models that were devised in order to calculate: (i) overall odds-ratios of acquiring the disease due to a specific risk factor by geographical region; (ii) overall odds-ratios of acquiring the disease from the different risk factors; (iii) overall risks of disease from consumption of ready-to-eat and barbecued foods; and (iv) overall effects of food handling (i.e., consuming food in raw, undercooked or unwashed state) and food preparation setting (i.e., eating food prepared outside the home) on risk of disease. The procedures for sensitivity analysis and removal of any influential and potentially-biased odds-ratio; and two methods for publication bias assessment are outlined. Finally, details are given on deviations from the standard risk categorisation scheme for specific foodborne hazards.

1. Introduction

Foodborne diseases caused by a variety of bacteria, parasites and viruses represent a growing public health issue worldwide. Annually, 1 out of 10 people in the world suffer from foodborne disease leading to a global burden of 33 (95% uncertainty interval of 25–46) million Disability-Adjusted Life Years (DALYs) (WHO, 2015). Diarrheal diseases cause more than half of global burden DALYs (WHO, 2015). Prevention of food-borne diseases is challenging and emphasised by the necessity to eliminate or reduce the contamination of food and food exports; the harm to human health and its associated costs to the health sector and economy; and the harm to livestock of economic importance. In order to design effective intervention strategies to reduce the burden of disease, an understanding of the risk factors for enteric infections is needed. However, these risk factors are wide-ranging and their identification is challenging.

Epidemiological studies make it possible to assess the risk associated with a particular source of exposure in the absence of randomised control trials. Among observational studies, case-control and transversal cohort designs offer the opportunity to evaluate the relevant

* Corresponding author.
E-mail address: ubarron@ipb.pt (U. Gonzales-Barron).

https://doi.org/10.1016/j.mran.2019.07.003
Received 13 April 2019; Received in revised form 21 July 2019; Accepted 22 July 2019
2352-3522/ © 2019 Elsevier B.V. All rights reserved.

statistical associations of potential risk factors with the occurrence of disease. Broadly speaking, a case-control study has been defined as an investigation into a relationship between a given disease and one or more potential causal or preventive factors, in which persons selected because they have the disease (i.e., cases) and selected persons who do not have the disease (i.e., controls) are compared in terms on their exposure to the factor under study. Case-control studies are considered to be an alternative to cohort studies, although they are clearly distinguished by two features: (i) sampling by disease in contrast to sampling by exposure, and (ii) investigative movement from-effect-to-cause in contrast to from-cause-to-effect (Kopec and Esdaile, 1990). In cohort studies, all potential cases are classified as exposed or unexposed. The odds ratio is the measure of association used in case-control studies. It is derived by dividing the odds of exposure among cases by the odds of exposure among controls.

The meta-analytical combination of odds-ratios, estimated from different observational studies of *sporadic* disease, produces a more precise estimate of the association between risk factor and disease, which leads to an increased statistical power. Furthermore, since case-control studies are performed using different populations, combining them through meta-analysis would produce odds-ratio estimates that have broader generalisability than is possible using only a single study. Such well-known strengths of meta-analysis are, nonetheless, contingent on the key condition that the measures share the same theoretical concepts and are of at least satisfactory validity (related to the principle of 'rubbish in, rubbish out'). Case-control and transversal cohort studies are regarded as being susceptible to bias, namely, selection bias, misclassification bias, recall bias and confounding, which can be introduced anytime in the design or conduct of the study, and can greatly impact the interpretation and validity of the study findings. Thus, observational studies that are potential candidates for meta-analysis should first undergo a systematic assessment of methodological quality.

The objective of this article is to describe, in detail, the systematic procedure established by the research group for (i) mapping all relevant literature reporting the results of primary observational studies on risk factors associated with eleven foodborne diseases; and (ii) combining the evidence from quality primary studies to estimate meta-analytical risks of acquiring foodborne diseases from exposure to different sources. The enteric diseases, for which systematic review and meta-analysis were conducted, were those caused by *Salmonella* non-typhi/paratyphi spp, *Campylobacter*, Shigatoxin-producing *Escherichia. coli* (STEC), *Listeria monocytogenes, Yersinia. enterocolitica, Toxoplasma gondii*, norovirus, Hepatitis A virus, Hepatitis E virus, *Cryptosporidium* spp. and *Giardia duodenalis*. Since primary studies investigate a variety of potential determinants of disease, a source categorisation scheme was developed for this meta-analysis study, which included travel, host-specific factors and hierarchical pathways of exposure – comprising person-to-person, animal, environment and food routes of transmission. In this way, the risk factors as labelled in the primary studies underwent, in as much as possible, a stratification that supported the integration and combination of results from all collected studies. In the framework of source attribution, the pathogens were selected based on the incidence of the foodborne illnesses in France (>1 million cases per year) and a multiplicity of potential sources (Van Cauteren et al., 2017).

Although meta-analyses of risk factors of foodborne pathogens are not new; they have been conducted for one pathogen at a time, for a selection of risk factors, and with specific inclusion criteria (Belluco et al., 2017; Boughattas, 2017; Bouzid et al., 2018; Domingues et al., 2012a,b; Kintz et al., 2017; Wei et al., 2016). Our proposal has taken a step forward, by formulating, as far as possible, a general approach that could be employed for all pathogens, yet conservative enough to keep most of the information provided in the primary studies. This had, in turn, some consequences for the statistical modelling. For instance, because most of the odds-ratios (and relative risks) were estimated with univariate analysis, they were retained in meta-analysis. However, as the dependence between studies can be related to the way the effect size is calculated (Van den Noortgate, 2013), this specific effect (type of analysis) had to be taken into account in the meta-analysis model. At the end, there had to be a trade-off between keeping most of the available information and removing non-satisfying or non-relevant information. The same kind of compromise had to be attained in the extent of aggregation of risk factors inside a category, as they are described in original papers. Is it pertinent to put within the same category lettuce and carrots, washed and not washed vegetables? If the level of aggregation was too strict, meta-analysis would not feasible, and if it was too large, the interpretation would be difficult. Those general methodological choices are described here.

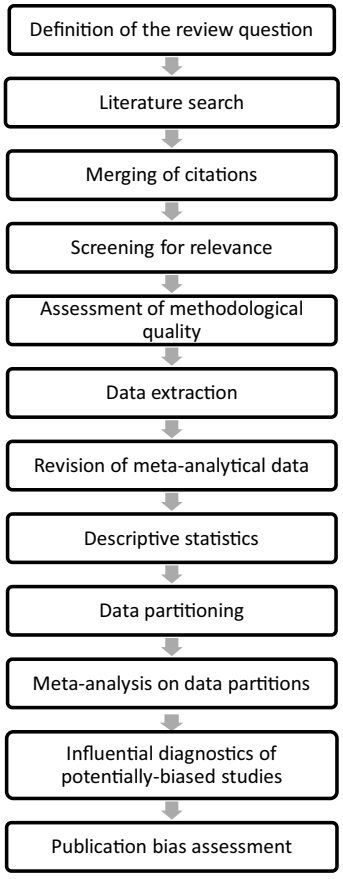

Fig. 1. Schematic of the steps followed for the systematic review and meta-analysis of observational studies of human enteric diseases.

2. Systematic review and meta-analysis

A methodical and reproducible procedure was planned and implemented for systematic review and meta-analysis of the sources of eleven foodborne diseases. Briefly (Fig. 1), for each of the eleven foodborne pathogens, meta-analysis began with a definition of the review question; clear enough to lead to a relevant set of keywords for literature search. Each reference record was screened for relevance for inclusion in the meta-analysis study, and subsequently, the methodological quality of the "candidate" studies were assessed using pre-set quality criteria. Those primary studies deemed as being possibly affected by any bias were marked as having "potential-for-bias". Previously-defined qualitative and quantitative information was then extracted from the observational studies. The joint meta-analytical data was first described using basic descriptive statistics. Next, data was partitioned into subsets of meaningful categories in order to make the most of the information contained in the (normally-sparse) meta-analytical data sets. Meta-analysis models were then fitted to each of the

data partitions or subsets. Once a meta-analysis model was fitted, influential diagnostics statistics were assessed in order to remove any influential observation originating from studies marked as having potential-for-bias. Publication bias was also evaluated. The steps are explained in depth in the following Sections.

2.1. Definition of the review question

Following the EFSA's guidelines for systematic reviews on food and feed safety assessments (EFSA, 2010), the review question was identified to have a typical PECO structure (viz. *population, exposure, comparator* and *outcome* as key elements) since it seeks to evaluate the association between a (risk) factor and disease in a population exposed to it. However, it must be born in mind that the fact that individual observational studies specify different risk factors makes the review question to be of a rather *open* nature. An open-framed question is a question that lacks some specification; in our case, the risk factors or pathways of transmission themselves. While most of the question's key elements can be identified, the *exposure* or potential risk factor is therefore study-specific; hence, the need for developing a harmonised risk categorisation scheme that could sustain the integration of primary results in this wide-ranging meta-analysis study. The question's key elements can be broken down, as follows:

Population: Cases of foodborne diseases in humans, which can be children, adults, elderly, pregnant women, immunocompromised or mixed population.
Comparator: Individuals free of disease or with another gastrointestinal disease (for instance, in a case-case study), which can be children, adults, elderly, pregnant women, immunocompromised or mixed population.
Outcome: A measure of association between exposure and disease, either odds-ratio (OR) or risk ratio (RR), when there is sufficient data to convert it into OR
Exposure: Risk factors of the specific foodborne disease.

The exposure is given by the possible determinants of foodborne human disease, which is precisely what the review question of this meta-analysis study seeks to assess. The risk factors were hierarchically categorised into travel, host-specific factors and pathways of exposure (i.e., person-to-person, animal, environment and food routes; Table 1). Each primary study investigated a specific set of risk factors or routes of transmission, which were attributed to defined categories at a later stage.

2.2. Literature search

Literature search for the 11 biological hazards was conducted in March 2017, and was limited to the languages English, French, Portuguese and Spanish. However, no restrictions were defined for the year of the study or type of publication. Relevant studies were identified from five bibliographic search engines, Science Direct, PubMed, Scielo, ISI Web of Science and Scopus. Searches were separately conducted for each foodborne hazard using a combination of keywords of the target *foodborne hazard, general terms* and *additional terms*, joined by the logical connector AND. Engines were set to search for these terms in title/abstract/keywords only.

Foodborne hazard or *disease*: (Salmonella OR salmonellosis) OR (Campylobacter OR campylobacteriosis) OR (Shiga* *Escherichia coli* OR Shiga* *E. coli* OR STEC OR VTEC OR EHEC OR O157:H7 OR O26:H11 OR O145:H28 OR O103:H2 OR O111:H8 OR O104:H4) OR (Listeria monocytogenes OR listeriosis) OR (*Yersinia enterocolitica* OR *yersiniosis*) OR (Toxoplasma OR toxoplasmosis) OR (Norovirus OR Norwalk) OR hepatitis A OR hepatitis E OR (Cryptosporidium OR crystosporidiosis) OR (Giardia OR giardiasis)

Table 1
Categorisation scheme of risk factors defined for transmission of foodborne diseases, with outer-to-inner hierarchies stored in the variables "Subcategory", "Route" and "Food Class".

Risk Factor	Subcategory	Route	Food Class
Travel	Any		
	Abroad		
	Inside		
Host-specific	Antiacids		
	Antibiotics		
	Chronic diseases		
	Other medical conditions		
	Breastfed		
	Poor personal hygiene*		
	Immunocompromising conditions*		
	Malnutrition*		
Pathway	Person to person	No further class or	
		Direct contact*	
		Community contact*	
		Sexual contagion*	
		Indirect contact*	
		Drug use*	
		Blood transfusion*	
		Occupational*	
	Animals	Farm animals	
		Pets	
		Wild animals	
		Occupational exposure	
	Environment	Day care attendance	
		Untreated drinking water	
		Recreational water	
		Farm/rural environment	
		Playground	
		Waste water*	
	Food	Meat	Beef
			Pork
			Other red meats
			Poultry
			Processed meat
			Others
		Egg	Eggs
			Egg products
		Dairy	Milk
			Cheese
			Fermented
			Fats
			Powder
			Undefined
		Produce	Fruits
			Vegetables
			Roots
			Legumes
			Sprouts
			Fungi
			Nuts
			Spices & Herbs
			Veg products
		Grains	Baked goods
			Cereals
			Pasta
			Others
		Seafood	Fish
			Crustaceans
			Molluscs
			Processed
			Undefined

(*continued on next page*)

ARTICLE IN PRESS

Table 1 (*continued*)

Risk Factor	Subcategory	Route	Food Class
		Beverages	Bottled water
			Alcoholic
			Soft drink
			Juices
			Undefined
		Composite	RTE composite
			Sandwich
			Dishes (including soups)
			Fast-food
		Sugars	No further class
		Seasoning	No further class

* Classes of host-specific factors and pathways of transmission used only for certain foodborne diseases (Refer to Section 3).

General terms: "case-control" OR "risk factor" OR "cohort"
Additional terms: infection OR disease

The term "sporadic" was not added because case-control studies conducted during outbreak investigations were also integrated into the reference databases. In addition to the references found by the bibliographic engines, other studies published as reports and conference proceedings were also searched. For every foodborne hazard, the records from the five search engines and those manually found were combined into a single reference database using the JabRef v. 3.0 software (Wassenhoven, 2010). Using this reference manager, duplicate references were carefully checked and deleted. Table 2 lists the number of citations recovered by bibliographic search engine and the total number of references after duplicate removal for each of the foodborne hazards under study. For most of the hazards, the search engines Web of Science and Scopus retrieved together more citations than Science Direct, PubMed and Scielo combined. Nonetheless, searches were still conducted in Science Direct, PubMed and Scielo because their results were mostly different from those of the broader engines. For instance, Scielo engine retrieved articles mostly written in Portuguese and Spanish carried out in Latin America.

2.3. Screening for relevance

After a reference database was built per foodborne pathogen, each reference from the 11 bibliographic databases was subjected to screening for its relevance to answering the review question in a two-step process. As a first step, a reference's title and abstract were appraised on the basis of two specific inclusion criteria: (i) disease in humans caused by any of the eleven foodborne hazards; and (ii) case-control studies conducted as part of *sporadic illness* or *outbreak* investigation; or alternatively cohort studies, excepting for *Salmonella* and *Campylobacter* reference databases. Subsequently, full-texts of the relevant records were appended to the references in the JabRef software, and read throughout by two researchers in order to assess their suitability for inclusion in the meta-analysis. The criteria for inclusion were: (i) information on case definition; (ii) information on statistical analysis, including results as OR measures, or sufficient information provided to calculate OR or to transform RR into OR; and (iii) assessment of relevant risk factors. Primary studies assessing "irrelevant" risk factors, such as those of demographic/socioeconomic nature (i.e., age, sex, parents' level of education, household income, type of dwelling, etc.) or confusing cause-and-consequence nature (i.e., stunting) were not included in the meta-analysis. Furthermore, reference lists from relevant papers were screened for additional eligible articles. Whenever found, "new" references were manually inserted in the appropriate database until December 2017.

Table 2
Number of records retrieved by search engine along with the total number of records without duplicates per reference database (foodborne hazard).

Reference database	Science Direct	PubMed	Scielo	Web of Science	Scopus	Total records, no duplicates
Salmonella	636	756	46	2456	5457	3858
Campylobacter	308	727	12	1141	2265	2360
Shiga toxin *E. coli*	978	951	460	2693	1113	4718
L. monocytogenes	1307	255	78	489	2399	1902
Y. enterocolitica	36	62	25	110	574	477
Toxoplasma	345	329	15	866	1210	1640
Cryptosporidium	211	207	51	1082	3539	1985
Giardia	61	69	29	352	788	691
Norovirus	98	96	21	365	457	672
Hepatitis A	176	268	127	153	1316	1624
Hepatitis E	61	100	127	58	429	614

2.4. Assessment of methodological quality

Through a standardised assessment, two researchers examined independently the methodological quality of each "relevant" primary case-control study utilising a checklist comprised of six areas of concern (Fullerton et al., 2012) (Fig. 2):

a Appropriate selection of the controls, in order to avoid *selection bias*;
b Adjustment to, at least partially, correct for confounders. The relevant confusion factors were collected;
c Strategy for the comparability between cases and controls: acceptable matching criteria for matched study designs (for instance, age and gender)
d Data analysis appropriate to the type of design, matched or unmatched;
e Acceptable responses rates for the exposed and control groups;
f Provision of crude and/or adjusted ORs and either confidence interval or p-value; alternatively, provision of sufficient data to calculate ORs;
g Overall quality of underlying methods, appropriate statistics, sensible data and quality of reporting/interpretation (combination of criteria a-f)

In the case of cohort studies, their quality was judged across the following areas of concern:

a Appropriate selection of cohorts;
b Comparability of cohorts: non-exposed cohort drawn from the same community as the exposed cohort;
c Provision of sufficient data to calculate ORs or convert RR to OR.
d Overall quality of underlying methods, ascertainment of exposure, sensible data and quality of reporting/interpretation (combination of criteria a-c)

For every primary study, each of the above items was answered. Non-compliance with a single criterion (items ***a*** to ***f*** for case-control studies, and items ***a*** to ***c*** for cohort studies) was not sufficient to reject a study. Instead, studies that failed to meet at least one of the checklist items (excepting item *g* for case-control studies and item *d* for cohort studies) were identified as "*having potential for bias*". Case-control or cohort studies that failed to meet item *g* or *d*, respectively, were discarded from the meta-analysis. The final decision of removing a primary study was reached by the consensus of scientific experts.

2.5. Data extraction

Data from primary studies that passed the previous steps were manually extracted by one trained reviewer using a standardised

ARTICLE IN PRESS

Study ID	
Study design	
	Case definition
	Are healthy controls used?
	Type of design Unmatched □ Random sampling □ Frequency matching □ Matched □ Individual matching □
	Matching criteria/Confounders:
	Target population:
	Case response rates acceptable?
	Control response rates acceptable?
	Are power and sample size calculations reported?
Exposure assessment	
	Was standard questionnaire used?
	Method of interview (phone, face-to-face, self-administered):
Analysis	
	Method of analysis Chi-square □ MH □ Logistic □ Conditional logistic □
	Is the method appropriate for the design?
	Are univariate Odd Ratios reported?
	Are they adjusted for confounders? Which?
	Are multivariate Odd Ratios reported?
	Are they adjusted for confounders? Which?
	Are PAF values reported?
Potential for bias	
	Quality of reporting
Overall	□ **Not sufficient information to judge = 1** □ **Evidence of bias = 1** Controls had another GI condition, inappropriate design/data analysis □ Standard rules of case-control studies = 0 □ Above standard rules of case-control studies = 0

Fig. 2. Methodological quality assessment form that informed the decision on the study's potential for having arrived to biased OR estimates.

spreadsheet. The information extracted was stored into the following 34 variables: country, year, year category, geographical region, serotype, phage type, population type, case definition, number of cases, number of exposed cases, median/mean age of cases, number of controls, number of exposed controls, median/mean age of controls, experimental design, type of model, type of analysis, type of OR, type of risk factor, stratification of the risk factor (comprising the outer-to-inner variables: subcategory, route and food class), RTE status, setting, handling, original label, OR value, OR lower and upper bounds, p-value of OR, matching criteria, adjusting criteria, potential-for-bias status, population attributable fraction (PAF) and PAF lower and upper bounds. For a very detailed explanation of each of the variables, see Supplementary Material I.

Two important variables were Case Definition and Design. A case was defined as “Lab-based” if the pathogen was isolated from a stool or blood culture for all of the cases; or “Either” if the diagnosis of some cases was lab-confirmed while others were diagnosed on the grounds of clinical features of the disease. This approach was conservative because, whichever the case definition, lab-based or either, the study was kept in the analysis. Although, the definition from the same pathogen can vary slightly between publications, and diagnosis tools can have different performance values, we consider keeping all the available data, and excluding only primary studies whose cases were all defined on the basis of symptoms. This problem, which ultimately is of different importance between pathogens, is introduced as a matter of discussion in each pathogen's paper.

The Design refers to the type of sampling of controls used in a case-control study. The design was defined as “Unmatched” when the control group was selected at random, or more often, when the case individuals were selected in such a way that their distribution of certain sociodemographic variable(s), for instance, age, was similar to that of the case group (also known, as stratification or frequency-matching). The design variable was set to “Matched” when controls were selected by case to match individual cases in certain specified respects (i.e., matching variables or confounders) so that one or more controls were paired to every case. A design of case-to-case matching was rarely reported, and albeit not excluded from the analysis, their results were given the potential-for-bias status. Once the data extraction spreadsheet from a foodborne hazard was finalised, a second researcher checked all the collected data against the original studies.

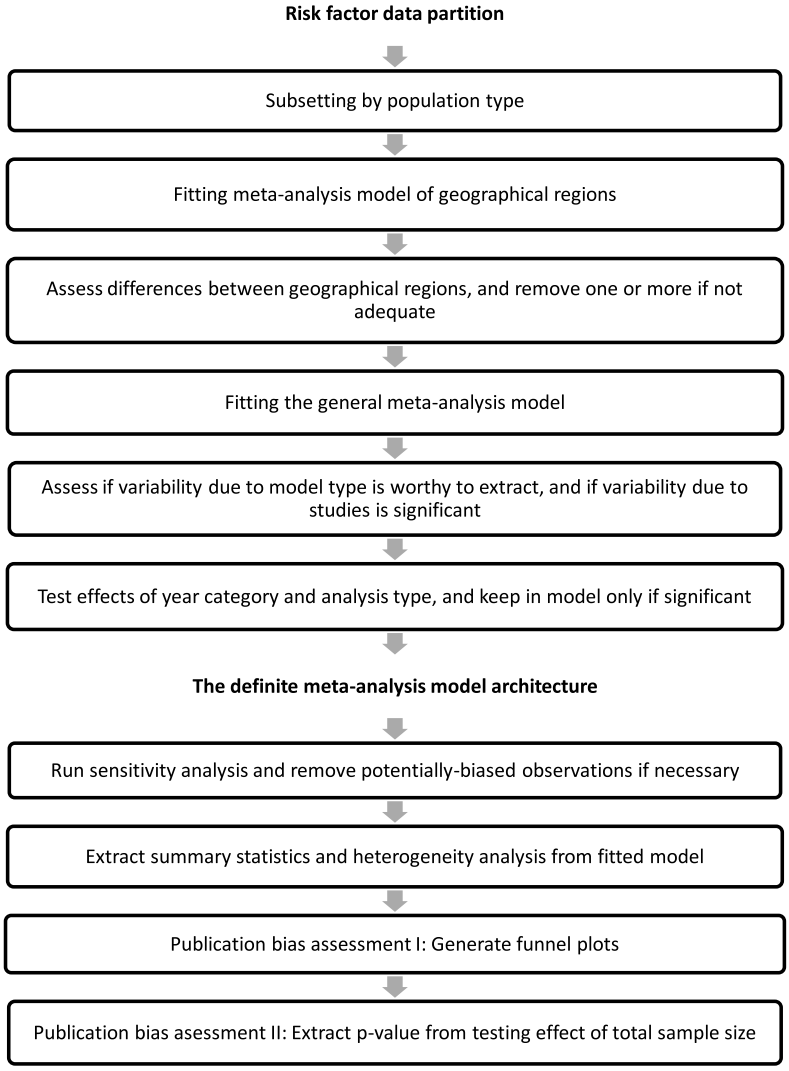

Fig. 3. Schematised sequence of the statistical analysis undertaken on every risk factor data partition. Boxes indicate actions taken, and absence of boxes represents stages.

2.6. Relevant rules for inclusion of ORs in the meta-analysis

a) In all cases except breastfeeding, the risk factors extracted from the primary studies were deemed as potential determinants of disease. Thus, when the results from a study were reported in the form of protective exposures, such as “drinking potable water as opposed to untreated water”, “handwashing after going to toilet” or “washing vegetables as opposed to non-washing vegetables”, the mean ORs and 95% CI annotated in the data extraction spreadsheet corresponded to the reciprocal of those values reported in the primary study. Accordingly, the exposure label was inverted (i.e., “drinking untreated water”, “no handwashing after going to toilet” and “non-washing vegetables”) while the number of exposed cases and exposed controls were adjusted.
b) Risk factors presenting a significant protective effect ($p < 0.05$), with no clear justification were not extracted. For example, ORs suggesting a protective effect of the consumption of vegetables against salmonellosis. Several explanations have been offered for this, including bias due to differential recall of exposure between case-patients and controls or due to confounding. Confounding could be at play if protective factors are in association with other unmeasured factors that affect risk, such as lifestyle factors or socioeconomic factors. It is also possible for “protective” associations to be observed with exposures that are actually risky, but that occur frequently and lead to protective immunity (Fullerton et al., 2012).
c) When primary studies investigated associations between risk factors and disease in both univariate and multivariate mode, both OR outcomes were recorded in the spreadsheet, assigning the class “Uni” or “Multi”, respectively, in the variable “Analysis Type” (See Supplementary Material I).

2.7. Synthesis of research on the risk factors for enteric disease transmission

The meta-analytical synthesis of the associations between transmission of enteric disease and risk factors was undertaken separately for every biological hazard. In general, the analyses sought to meet four objectives:

a) Descriptive statistics of the meta-analytical data on risk factors of disease;
b) Calculation of the overall risks of acquiring the disease due to travel, host-specific factors and transmission pathways related to person-to-person contagion, animal contact, environmental exposures and food vehicles;
c) Calculation of the overall risks of disease from consumption of ready-to-eat (RTE) and barbecued foods (BBQ); and

Table 3
Interrelationships between type of analysis, models and OR type considered for the meta-analytical data.

Analysis type	Model	OR type
Univariate analysis	Chi-square	Raw
	Mantel-Haenzel	Raw or adjusted
	Unconditional logistic	Raw or adjusted
	Conditional logistic	Raw or adjusted
Multivariate analysis	Unconditional logistic	Adjusted
	Conditional logistic	Adjusted

d) Calculation of the overall effects of food handling and food preparation setting on risk of disease.

However, some biological hazards did not undergo the five types of analysis, due to either data sparseness, lack of data or irrelevance of the analysis for the particular hazard. For instance, analysis (c) was not carried out for *Cryptosporidium* spp, since the transmission of human cryptosporidiosis through consumption of RTE or BBQ foods is meaningless, and therefore such association was not investigated in the primary studies. All meta-analysis modelling as well as sensitivity analysis, publication bias analysis and meta-analytical graphs were executed in the R software (R Development Core Team, 2008) implemented with the metafor package (Viechtbauer, 2010). The meta-analysis procedures for each of the four research synthesis objectives are described in depth, as follows.

2.7.1. Descriptive statistics

For every foodborne disease, a PRISMA diagram showing the detailed results of the study selection process was prepared as well as a summary table containing the most relevant characteristics of the observational studies selected for meta-analysis. In addition, for every foodborne hazard, the descriptive statistics of the meta-analytical data set included range of study's publication year; number of ORs by country, population type, serovar, experimental design, risk factor, analysis type, model type, OR type; and number of publications with potential bias.

2.7.2. Meta-analysis for host-specific risk factors and travel-, person-to-person environment-, animal- and food-related pathways of transmission

Next, the overall risk of foodborne illness transmission for every exposure or source was calculated through systematic meta-analyses on the odds-ratio parameterisation. To obtain the overall OR for every risk factor, the meta-analytical data was first partitioned into the risk factor subsets presented in Table 1. The sequence of statistical analyses carried out on every risk factor partition was systematic, and is schematised in Fig. 3. First, a *meta-analysis model by region* was adjusted to a data partition in order to summarise the associations between exposure and illness (i.e. pooled OR) for every geographical region (r; Eq. (1))

$$lnOR_{ir} = \beta_{0i} + \beta_{1r} Region_r + \varepsilon_{ir}$$
$$\beta_{0i} = \overline{\beta_0} + u_i \quad (1)$$

To extract the between-study variability, the model's intercept β_{0i} was allowed to shift according to the random effect u_i due to the deviation of the primary study i from the mean estimate $\overline{\beta_0}$. The random effects u_i were assumed to be normally-distributed with mean zero and variance s_u^2, while the errors ε_{ir} assumed to follow a normal distribution with mean zero and variance s^2. By setting the intercept mean to zero, the pooled ln(OR) values for the geographical regions available in the data partition were obtained in the vector β_{1r}. A geographical region class consisting of less than 3 ORs was removed from the analysis. The objective of this region-specific meta-analysis was to inform the decision on the geographical regions that should be kept for the subsequent pooling of ORs.

Second, the *general meta-analysis model*,

$$lnOR_{ikt(j)} = \beta_{ik} Subgroup_k + u_{t(j)} + \varepsilon_{ikt(j)}$$
$$\beta_{ik} = \bar{\beta}_k + v_{ik} \quad (2)$$

was fitted to every data partition,where k is a subgroup class that depends on data partition. For instance, if the data partition is Travel, the variable Subgroup consists of the classes: Inside, Abroad and Any ($k = 1, 2, 3$). Similarly, if the data partition is Eggs, the subgroup is comprised of the classes Eggs and Egg Products ($k = 1, 2$). Thus, bear in mind that the levels of the variable Subgroup are not anything else but risk factors. The primary study is denoted by i, while the type of analysis, multivariate or univariate, is denoted by t. For multivariate analysis ($t = 1$), model type j takes two levels (unconditional logistic and conditional logistic) whereas for univariate analysis ($t = 2$), model type (or estimation method) j takes four levels (chi-square, Mantel-Haenszel, unconditional logistic and conditional logistic) (Table 3). It is expected that the type of model used to compute ln OR, either in univariate or multivariate analysis, is responsible for some of the heterogeneity in the OR values retrieved from the literature for a given exposure. To extract this variability from the overall estimates, random effects $u_{t(j)}$ with model j nested within analysis type t as subject of variation were placed on the intercept (which was set to a mean of zero). These intercept nested random effects $u_{t(j)}$ were assumed to follow a normal distribution with mean zero and variance s_u^2.

To extract the between-study heterogeneity from each of the k subgroup levels (viz. categorised risk factors), random effects v_{i1}, v_{i2}, ...v_{ik} having primary study i as subject of variation were placed on the subgroup-level coefficients β_1, β_2, ...β_k. These subgroup-level random effects v_{i1}, v_{i2}, ...v_{ik} were assumed to be normally-distributed with mean zero and uncorrelated variance s^2_{v1}, s^2_{v2}, ... s^2_{vk}, respectively. In this way, the fitted coefficients $\bar{\beta}_1$, $\bar{\beta}_2$, ...$\bar{\beta}_k$ represent the overall (or meta-analysis estimate) ln(OR) for the subgroup levels 1, 2, .. k, respectively; whereas the variances s^2_{v1}, s^2_{v2}, ... s^2_{vk} represented the between-study variances (widely known as τ^2 in meta-analysis) corresponding to the subgroup levels or exposures. The errors or residuals $\varepsilon_{ikt(j)}$ were also assumed to follow a normal distribution with mean zero and variance s^2.

Every time a general meta-analysis model (Eq. (2)) was adjusted to a data partition, the significance of the variances due to model type and primary study were assessed. If the variance associated to model type was not significant, the nested random effects were removed from the general model, and instead a fixed-effects term for AnalysisType was added in Eq. (2). If the AnalysisType coefficient was significant ($p < 0.10$), it was kept in the model. However, in cases where data points were too few in at least one of the classes of analysis type (for instance, very few multivariate ORs in contrast to many univariate ORs), the model was further abridged by removing the AnalysisType fixed-effects altogether. The random effects v_{ik} for extracting the between-study variability of subgroups (i.e., risk factors), were maintained in all data partitions for which a random-effects model was the most suitable meta-analysis solution. In cases where the between-study variance turned out to be non-significant or zero, the random effects v_{ik} were dropped from the model. In this way, meta-analysis models could be further simplified to a purely fixed-effects solution if the variances due to model type and due to primary study were both non-significant. This was very likely to occur in partitions consisting of very few data points.

Once the best random- and/or fixed-effects structure, accounting for variability due to model type and primary study, was defined (as explained above); the effect of year category (i.e., before 2000 and after 2000) on the ORs retrieved from the literature was appraised by linearly adding the covariate YearCat to the fixed-effects linear specification,

$$lnOR_{ikt(j)} = \beta_{ik} Subgroup_k + \beta_y YearCat_y + u_{t(j)} + \varepsilon_{ikt(j)}$$
$$\beta_{ik} = \bar{\beta}_k + v_{ik} \quad (3)$$

The moderator year category was dropped from the model if its associated p-value was higher than 0.15. Else, it remained in the model and the overall ORs of the subgroup levels, β_{ik}, were estimated with basis on the more recent primary studies (year category after 2000). Likewise, when the covariate AnalysisType was significant, the overall ORs of the subgroup levels were expressed on multivariate "scale".

2.7.3. Meta-analysis on the risk of gastrointestinal illness transmission by consumption of ready-to-eat (RTE) foods and barbecued (BBQ) foods

For some microbiological hazards, two additional meta-analysis models were adjusted in order to estimate the overall risk of transmission of illness through the consumption of RTE foods and BBQ foods. Two data partitions were created, one that gathered all data for which the "RTE" variable status was set to "1″, and another that brought together all data whose labels contained the word "BBQ". Due to the fewer data points contained in these partitions, meta-analytical models were simplified, according to Eq. (4),

$$\ln OR_i = \beta_{0i} + \varepsilon_i$$
$$\beta_{0i} = \bar{\beta}_0 + u_i \quad (4)$$

where $\bar{\beta}_0$ is the overall ln(OR) for consumption of RTE foods or BBQ foods, accordingly. The subject of variation of the intercept random effects u_i was primary study: and u_i were assumed to follow a normal distribution with mean zero and variance s_u^2.The residuals ε_i were also assumed to follow a normal distribution with mean zero and variance s^2. Model (4) was adjusted separately to the RTE and BBQ data partitions.

The meta-analytical models shown in Eqs. (1)–(4) were fitted separately by population type; namely, *mixed, children* and *susceptible*. Thus, a further partitioning of data by population type was further carried out before model fitting. However, when data points from the children or susceptible population were too few to justify a separate meta-analysis, its data was merged with that of the mixed population. For all of the hazards, data from the adult population was integrated to that of the mixed population.

2.7.4. Meta-analysis on the effects of handling and setting on the risk of foodborne illness transmission by food consumption

The second group of meta-analyses only pertained to food vehicles; and their purpose was to estimate, for some food classes, how the overall OR (hereafter referred to as *base OR*) was affected by mishandling of food (i.e., eating raw, eating undercooked or eating unwashed food) or by eating food prepared outside the home (i.e., eating out as setting). To obtain overall estimates of the effects of *Handling* (m = raw, undercooked or unwashed state) and *Setting* (n = home, eating out) on the risk of contracting an enteric illness, the general meta-analysis model,

$$\ln OR_{itmn} = \beta_{0i} + \beta_1 AnalysisType_t + \beta_{2m} Handling_m + \beta_{3n} Setting_n + \varepsilon_{itmn}$$
$$\beta_{0i} = \bar{\beta}_0 + u_i \quad (5)$$

was fitted to food-class data partitions conveniently selected for containing sufficient data across the Handling or Setting levels. Notice that, although the variable "Handling" consists of five levels, namely, "Raw", "Undercooked", "Unwashed", "Poor" and "Preparation" (Supplementary Material I), only the first three levels were used for this meta-analysis. The effect of analysis type β_1 was tested since multivariate ORs tended to be higher than univariate ORs. AnalysisType only remained in the model if its corresponding p-value was < 0.10. Random effects u_i were placed only on the intercept β_0 to withdraw the between-study variability s_u^2. Deviations u_i were assumed to follow a normal distribution with mean zero and variance s_u^2. The residuals ε_{itmn} were also assumed to follow a normal distribution with mean zero and variance s^2

The exponential of the parameters β_{2m} and β_{3n} can be interpreted as the average number of times mishandling food or eating out, respectively, increased the base OR. This is because, in mathematical terms, the ratio of the mean OR when food is mishandled (or, alternatively, when food is prepared outside the home) to the base OR is defined by the estimate "$\frac{\exp(\beta_0+\beta_{2m})}{\exp(\beta_0)} = \exp(\beta_{2m})$" for handling or "$\frac{\exp(\beta_0+\beta_{3n})}{\exp(\beta_0)} = \exp(\beta_{3n})$" for setting. Examples of selected food classes in which this type of meta-analysis was conducted are: pork, poultry or processed meats to estimate the effects of eating undercooked or prepared outside the home; beef to estimate the effects of eating raw or undercooked; eggs to estimate the effect of eating raw or undercooked; vegetables to estimate the effect of not washing them; and milk/cheese to estimate the effect of consuming them in raw state. As an example, the meta-analysis on handling allowed to estimate that people who consumed vegetables *without previous proper washing* had on average 1.5 (95% CI: 1.0 – 2.2) times the odds of becoming infected with salmonellosis than those who *simply* consumed vegetables (which could include washed or not). Another way of interpreting the meta-analytical effect of handling or setting could be: People who consumed the food under question and acquired the disease were, on average $\exp(\beta_{2m})$ *or* $\exp(\beta_{3n})$ times more likely to have mishandled it or eaten out, respectively, than those people who consumed the same food yet did not become ill.

All meta-analysis models Eqs. (1)–(5) were essentially weighted random-effects linear regression models with weights defined as the inverse of the sampling variance of the reported ORs. In all random-effects models, the between-study variance was computed by the maximum likelihood (ML) estimator. The restrictive maximum likelihood (REML) estimator was not used because partitions were sparse, and previous simulation studies suggested that, under data sparseness, the REML method can underestimate the between-study variance (s_u^2), and, furthermore, its solution is often associated to a greater mean square error (Angeliki Veroniki et al., 2016). Wald-type confidence intervals were calculated under the assumption of normality.

2.8. Influential diagnostics of the risk measures with potential for bias

For each of the meta-analysis conducted on a data partition, a sensitivity analysis based on the Cook's distance approach was used to assess the influence of a data point on the meta-analysis estimates. An observation was considered as influential when its Cook's distance was greater than four times the mean Cook's distance. The validity of an influential point was then examined by checking its "potential for bias" status ("0" or "1"). If that particular OR measure was flagged as "1" (i.e., potentially biased), the observation was removed, and the meta-analysis model was re-adjusted. Fig. 4 shows an example of the procedure described above for the meta-analysis conducted on the risk

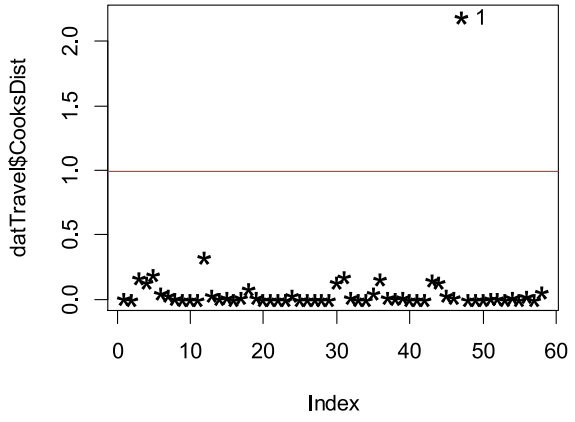

Fig. 4. An example of the use of Cook's distance as decision criterion for removal of potentially-biased odds ratios from the data sets.

factors of salmonellosis related to travel. Notice that, in this case, the observation #47 was potentially-biased (status "1" next to the marker), and therefore it was removed from the Travel data partition. Essentially, it is a conservative approach, because ORs are inspected one by one and only excluded if they come from potentially-biased results. Exceptionally, when forest-plots showed a study whose 95% CI overlapped only slightly with other studies, a re-analysis excluding this observation was carried out to evaluate the robustness of the meta-analysis. If the combined result changed significantly, the observation was excluded; and oppositely, if the combined result did not change substantially, the observation was kept (Wei et al., 2016).

2.9. Assessment of publication bias

In meta-analysis, it is widely known that if studies with small or non-significant results remain unpublished, the plot of the effect sizes and their standard errors produce an asymmetric funnel plot. One may be able to detect such asymmetry by testing whether the observed outcomes (or residuals from a model with moderators) are related to their corresponding sampling variances, standard errors, or more simply, sample sizes Viechtbauer, 2010). Thus, for the meta-analysis of every data partition, publication bias was first graphically assessed by the funnel plot of the model's residuals and their corresponding inverse standard errors. Next, the presence of publication bias was formally tested by including total sample size (total sample size N was defined as the sum of the number of cases and controls) as a covariate in the meta-analysis regression models. So, in a separate regression, the total sample size N was added to Eqs. (2)–(4) as an additional covariate. If the p-value associated to total sample size was non-significant (i.e., no effect of study size on the ORs, as measured in the primary studies), we concluded that the existence of a file-drawer problem was unlikely (Hox and de Leeuw, 2003) for the pathways or risk factors pertaining to the data partition being meta-analysed. A significant p-value for N implies that the OR value measured by the researchers depends upon the sample size; in other words, the tested level of association with disease is conditional to the study size. In this case, it is likely that small-sized studies have remained unpublished because of their failure to detect significant ORs.

2.10. General interpretation of meta-analysis results

For every gastrointestinal disease and every data partition, the meta-analysed risk factors are presented in summary tables only when significant. Pooled ORs were considered as significant when the lower bound of the 95% CI was equal or greater than 1 (with precision of 3 digits in decimal value), except for breastfeeding when the upper bound of the confidence interval should be below 1.

For transparency, apart from the significant pooled ORs and their confidence intervals, the summary tables of meta-analysis results provide details on the number of primary studies and the number of ORs used to estimate the pooled OR, the p-value of the publication bias test (i.e., testing the effect of study size, N), the between-study variability (τ^2), the QE test result for presence of residual heterogeneity, the variance of residuals (s^2) and the intra-class correlation (I^2). When the number of ORs/studies is small or the publication bias is significant, the interpretation of the pooled OR is done more cautiously. A significant QE statistic indicates the presence of residual between-study heterogeneity that could not be fully explained by the moderator(s) already considered in the meta-analysis model (Viechtbauer, 2010). It is worth mentioning that, in partitions where risk factors were not further broken down in classes (variable Subgroup devoid of meaning and hence dropped from Eq. (3)), the meta-analysis summary results do not present a QE statistic but a Q statistic, whose significance indicates the presence of heterogeneity between studies. The intra-class correlation (I^2) is the percentage of the total variance that is explained by the variation between the studies. Although the interpretation of I^2 is relative, the general rule suggests: 0% no heterogeneity, 25% low heterogeneity, 50% moderate heterogeneity, 75% high heterogeneity (Higgins et al., 2003). The number of points removed is also given, but because this sensitivity analysis result comes from a small number of potentially-biased studies and influential points, we assumed that the interpretation is not affected by this removal.

Meta-analytical forest plots were constructed for all risk factors relevant to the foodborne disease under study. The origin of the dataset as well as the relevance of the extent of aggregation inside a risk factor category were checked by a group of experts of each pathogen, in order to ensure the adequacy of the analysis and interpretation of the data; in particular when results were not completely expected. Other sources of information, such as natural pathways of contamination and recorded outbreaks, were also taken into account to interpret the causality between risk factor and disease, when it turned out to be statistically significant on meta-analysis. Supplementary material is provided in the meta-analysis study of every pathogen, containing a descriptive analysis of the pertaining meta-analytical dataset; funnel plots for all data partitions as a complementary tool to appraise publication bias; and some forest plots.

3. Pathogen-specific deviations from the standard data categorisation scheme

The data categorisation scheme described in depth in the Supplementary Material was implemented for the stratification of the information extracted for the eleven foodborne diseases. Yet, to attend to epidemiological aspects that are pathogen-specific, there were certain variations introduced in the standard categorisation scheme. The variable "**Serotype**", defined to accommodate the serotypes of pathogens such as *Salmonella* spp., *Y. enterocolitica* or *L. monocytogenes*, was also utilised to store the pathotypes for *E. coli*, assemblages for *G. duodenalis*, species for *Campylobacter* and *Cryptosporidium* and genogroup for norovirus. The variable "**PhageType**" was also used to record the strain of STEC.

For salmonellosis, campylobacteriosis, STEC infection, yersiniosis, cryptosporidiosis, giardiasis, norovirus infection, hepatitis A and hepatitis E, the variable "**Population**" was stratified into mixed, adult, children and susceptible. Nonetheless, within the susceptible category, distinctions were made for some diseases: (i) for toxoplasmosis, the susceptible population consisted only of pregnant women (mentally ill, and immunocompromised individuals was excluded, see specific paper in this Special Issue); (ii) for hepatitis E, pregnant population was separated from other susceptible; and (iii) for cryptosporidiosis, the susceptible population consisted only of HIV positive individuals. Only for STEC infection, the children population was subdivided into young (younger than 6 years) and older (older than 6 years) children. Listeriosis employed an altogether different stratification of population, comprising the classes "non-perinatal", "perinatal" and "peri/non-perinatal" (i.e., combination of both, since it was not possible to disaggregate them as per the design of the primary study).

The categorisation scheme for **host-specific** risk factors, consisting of five standard categories (Antiacids, Antibiotics, Chronic, OtherMed, Breastfed; Table 1), were employed for all diseases, except for giardiasis and norovirus infection. The host-specific risk factors considered for norovirus infection were poor personal hygiene (i.e., "Hygiene", with exposures such as: no handwashing after toilet, keeping poor hygiene habits, etc.) and suffering from an immunocompromising condition (i.e., "Immuno", with exposures such as organ transplant, HIV positive, etc.). For giardiasis, the same two classes explained above were used plus a third class, "Malnutrition", which was specific to malnourished children. In addition to the standard host-specific risk factors, the classes "Hygiene" and "Immuno" were also used for cryptosporidiosis, toxoplasmosis, hepatitis A and hepatitis E. In diseases where the host-specific risk factors included "Chronic", "Immuno" and "OtherMed" (viz. cryspostoridiosis, hepatitis A and hepatitis E), a rule was

established for classifying chronic immunological conditions within the "Immuno" class instead of the "Chronic" class. Hence, "Chronic" was left for the medical conditions that did not compromise the immunological system, while "OtherMed" class was always left for the unclassifiable conditions.

Within the pathways of transmission, "**WasteWater**" (i.e., contact with human fecal matter, lack of sanitation, open sewage, etc.) was an environmental-pathway class relevant to yersiniosis, giardiasis, cryptosporidiosis, toxoplasmosis, norovirus, hepatitis A and hepatitis E. **Person-to-person contagion** was a pathway of transmission without further subdivision for most of the enteric diseases. However for pathogens where person-to-person is an important way of transmission, such as cryptosporidiosis, norovirus, hepatitis A and hepatitis E, further subdivisions or classes were created when data was sufficient to do so. Person-to-person contagion implied exposures such as living in a household with ill person. For cryptosporidiosis, person-to-person transmission was stratified in three classes: "Contact" (contact in household), "ContactC" (contact in the community) and "Sex" (sexual transmission). In addition to these three classes, contagion of hepatitis A was considered to be also due to "IndContact" (indirect contact), "Drugs" (drugs injection) and "Transfusion" (contagion through blood transfusion). The strata for person-to-person contagion of hepatitis E were "Contact", "Sex", "Drugs" and "Occupational" (occupational contact with ill people). For toxoplasmosis, all of the person-to-person contagion due to sexual transmission was excluded from the analysis (see specific article in this Special Issue). Some other exceptional partitions were created when the question was epidemiologically relevant to the pathogen under study, and those are explained in the pertaining article. For instance, the cat exposure risk factor was analysed specifically for *Toxoplasma*, while the risk factor of consuming pig liver products was considered only for hepatitis E.

4. Concluding remarks

This broad meta-analysis of case-control studies allowed mapping, identification, categorisation and summarisation of risk factors for eleven foodborne diseases. Applying the unified methodology proposed in this article, it was found that risk factors such as consumption of undercooked beef (pooled OR = 2.052; 95% CI: 1.576 – 2.672), and poor handling of foods (pooled OR = 2.00; 95% CI: 1.958 – 2.504) were associated with toxoplasmosis in pregnant women (Thébault et al., 2019a); while top pathways of transmission of hepatitis E with serological definition in the mixed population turned out to be consumption of pork (pooled OR = 2.267; 95% CI: 1.675 – 3.068), contact with farm animals (pooled OR = 2.071; 95% CI: 1.506 – 2.848), blood transfusion (pooled OR = 1.732; 95% CI: 1.352 – 2.218) and drinking water (pooled OR = 1.692; 95% CI: 1.434 – 1.996) (Pavio et al., 2019). Risk factors of transmission of giardiasis in children were person-to-person contagion (pooled OR = 3.404; 95% CI: 1.873 – 6.187), consumption or contact with contaminated water (pooled OR from 1.863 to 2.059), and consumption of produce (pooled OR = 2.192; 95% CI: 1.465 – 3.278), while breastfeeding was found to exert a protective effect against acquiring giardiasis (pooled OR = 0.638; 95% CI: 0.499 – 0.817) (Thébault et al., 2019b). For campylobacteriosis in the mixed population, the highest associations were found for travel abroad (pooled OR = 4.626; 95% CI: 3.522 – 6.075), occupational exposure to animals/carcasses (pooled OR = 3.022; 95% CI: 2.264 – 4.036); recent use of anti-acids (pooled OR = 2.911; 95% CI: 2.040 – 4.154), contact with farm animals (pooled OR = 2.235; 95% CI: 1.891 – 2.641), consumption of meat of non-specified origin (pooled OR = 1.933; 95% CI: 1.695 – 2.204), and consumption of raw milk (pooled OR = 1.828; 95% CI: 1.313 – 2.545) (Fravalo et al., 2019).

The data categorisation scheme proposed and employed in the meta-analysis of risk factors for sporadic disease is original since it provides a unified framework for different water- and foodborne pathogens. Nonetheless, such a unified approach was also flexible enough to support pathogen-specific fine-tuning variations in data categorisation, which were introduced to provide more precise answers to epidemiological questions pertinent of the disease under study. Secondly, the meta-analysis modelling methodology takes into account the type of analysis and model as a source of heterogeneity, whenever data was sufficient. Across risk factor categories, the number of extracted ORs was very different between pathogens, making the statistical analysis more limited when the available data was too low. Furthermore, certain ways of transmission, of specific importance to the pathogen under study, called for different data partitioning. Furthermore, the heterogeneity of labelling and their meaning in a specific context could impinge on the accuracy of the risk categorisation process, so labels had to be interpreted specifically for each pathogen.

Supplementary Material I. Description of information extracted from observational studies

Declaration of Competing Interest

I hereby declare there is no conflict of interest related to the manuscript "Strategy for Systematic Review of Observational Studies and Meta-Analysis Modelling of Risk Factors for Sporadic Foodborne Diseases".

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mran.2019.07.003.

References

Angeliki Veroniki, A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J.P.T., Langan, D., Salanti, G., 2016. Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res. Synth. Methods 7, 55–79.

Belluco, S., Simonato, G., Mancin, M., Pietrobelli, M., Ricci, A., 2017. *Toxoplasma* gondii infection and food consumption: a systematic review and meta-analysis of case-controlled studies. Crit. Rev. Food Sci. Nutr. 11, 1–12.

Boughattas, S., 2017. *Toxoplasma* infection and milk consumption: meta-analysis of assumptions and evidences. Crit. Rev. Food Sci. Nutr. 57 (13), 2924–2933.

Bouzid, M., Kintz, E., Hunter, P.R., 2018. Risk factors for *Cryptosporidium* infection in low and middle income countries: a systematic review and meta-analysis. PLoS Negl. Trop. Dis. 12 (6), e0006553.

Domingues, A.R., Pires, S.M., Halasa, T., Hald, T., 2012a. Source attribution of human campylobacteriosis using a meta-analysis of case-control studies of sporadic infection. Epidemiol. Infect. 140 (6), 970–981.

Domingues, A.R., Pires, S.M., Halasa, T., Hald, T., 2012b. Source attribution of human salmonellosis using a meta-analysis of case-control studies of sporadic infection. Epidemiol. Infect. 140 (6), 959–969.

EFSA, 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. Eur. Food Saf. Auth. EFSA J. 8 (6), 1637. https://doi.org/10.2903/j.efsa.2010.1637. [90 pp.].

Fravalo, P., Kooh, P., David, J., Mughini-Gras, L., Cadavez, V., Gonzales-Barron, U., 2019. Meta-analysis on risk factors for sporadic campylobacteriosis. Microb. Risk Anal (in this issue).

Fullerton, K.E., Scallan, E., Kirk, M.D., Mahon, B.E., Angulo, F.J., de Valk, H., van Pelt, W., Gauci, C., Hauri, A.M., Majowicz, S., O'Brien, S., 2012. Case-control studies of sporadic enteric infections: a review and discussion of studies conducted internationally from 1990 to 2009. Foodborne Pathogen Dis. 9 (4), 281–292.

Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ 327 (7414), 557–560.

Hox, J.J., De Leeuw, E., 2003. Multilevel models for meta-analysis. In: Reise, S.P., Duan, N. (Eds.), Multilevel modelling: Methodological advances, Issues and Applications. Lawrence Erlbaum Associates, New Jersey, pp. 90–111.

Kintz, E., Brainard, J., Hooper, L., Hunter, P., 2017. Transmission pathways for sporadic shiga-toxin producing E. coli infections: a systematic review and meta-analysis. Int. J. Hyg. Environ. Health 220, 57–67.

Kopec, J.A., Esdaile, J.M., 1990. Bias in case-control studies. a review. J. Epidemiology Community Health 44 (3), 179–186.

Pavio, N., Thébault, A., Cadavez, V., Gonzales-Barron, U., 2019. Meta-analysis on risk factors for sporadic Hepatitis E infection. Microb. Risk Anal (in this issue).

R Development Core Team, 2008. R: A language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL. http://www.R-project.org.

Thébault, A., Kooh, P., Cadavez, V., Gonzales-Barron, U., Villena, I., 2019a. Meta-analysis on risk factors for sporadic toxoplasmosis. Microb. Risk Anal (in this issue).

Thébault, A., Jourdan-Da-Silva, N., Cadavez, V., Gonzales-Barron, U., Villena, I., 2019b.

ARTICLE IN PRESS

Meta-analysis on risk factors for sporadic giardiasis. Microb. Risk Anal.(in this issue).
Van Cauteren, D., Le Strat, Y., Sommen, C., Bruyand, M., Tourdjman, M., Jourdan Da Silva, N., Couturier, E., Fournet, N., de Valk, H., Desenclos, J.-.C., 2017. Estimated annual numbers of foodborne pathogen-associated illnesses, hospitalizations and deaths, France, 2008-2013. Emerg. Infect. Dis. 23 (9), 1486–1492.
Van den Noortgate, W., López-López, J.A., Marín-Martínez, F., Sánchez-Meca, J., 2013. Three-level meta-analysis of dependent effect sizes. Behav. Res. Method. 45 (2), 576–594.
Viechtbauer, W., 2010. Conducting meta-analyses in r with the metafor package. J. Stat. Softw. 36 (3), 1–48. Available from: http://www.jstatsoft.org/v36/i03/.
Wassenhoven, D., JabRef 2.4 Benutzerhandbuch, 2010, 77 p. (In German). Available from: https://manuals.jabref.org/JabRef-UserMan.
Wei, H.X., He, C., Yang, P.L., Lindsay, D.S., Peng, H.J., 2016. Relationship between cat contact and infection by *Toxoplasma Gondii* in humans: a meta-analysis. Comp. Parasitol. 83 (1), 11–19.