Risk Analysis, Vol. 29, No. 7, 2009 DOI: 10.1111/j.1539-6924.2009.01224.x

Source Attribution of Food-Borne Zoonoses in New Zealand: A Modified Hald Model

Petra Mullner,[1,2,*] Geoff Jones,[3] Alasdair Noble,[3] Simon E. F. Spencer,[1,4] Steve Hathaway,[2] and Nigel Peter French[1]

A Bayesian approach was developed by Hald *et al*.[1] to estimate the contribution of different food sources to the burden of human salmonellosis in Denmark. This article describes the development of several modifications that can be used to adapt the model to different countries and pathogens. Our modified Hald model has several advantages over the original approach, which include the introduction of uncertainty in the estimates of source prevalence and an improved strategy for identifiability. We have applied our modified model to the two major food-borne zoonoses in New Zealand, namely, campylobacteriosis and salmonellosis. Major challenges were the data quality for salmonellosis and the inclusion of environmental sources of campylobacteriosis. We conclude that by modifying the Hald model we have improved its identifiability, made it more applicable to countries with less intensive surveillance, and feasible for other pathogens, in particular with respect to the inclusion of nonfood sources. The wider application and better understanding of this approach is of particular importance due to the value of the model for decision making and risk management.

KEY WORDS: *Campylobacter*; food-borne zoonoses; microbial risk assessment; risk attribution; *Salmonella*

1. INTRODUCTION

Source attribution is the process of determining what proportion of a particular disease is acquired from a given source (e.g., poultry) and through a given pathway (e.g., food, water, and person-to-person transmission). This capacity to attribute cases of human disease to a food vehicle or another source responsible for illness is critical for the identification and prioritization of food safety interventions, and a variety of approaches are used worldwide.[2] Most quantitative risk assessments commonly deal with one pathogen occurring in a single food commodity and are targeted to identify options for prevention, intervention, and control. In contrast, source attribution models provide information about the public health impact of all important sources and pathways.[3] These methods are intended to provide decisionmakers with a set of tools for priority setting to achieve a more targeted control of diseases.[3] Overall, these tools are now vital components for the prioritization of hazards and interventions in food systems.[4]

Classical approaches to source attribution include full risk assessments, analysis and extrapolation of surveillance or outbreak data, and analytical epidemiological studies.[5–8] However,

[1] Molecular Epidemiology and Veterinary Public Health, Hopkirk Institute, Massey University, Palmerston North, New Zealand.
[2] New Zealand Food Safety Authority, Wellington, New Zealand.
[3] Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand.
[4] Current affiliation: School of Mathematical Sciences, University of Nottingham, Nottingham, UK.
* Address correspondence to Petra Mullner, Massey University, Palmerston North 4442, New Zealand; tel: +64(6)350-5270; fax: +64(6)350-5716; p.mullner@massey.ac.nz, petra.mullner@nzfsa.govt.nz.

recently, source attribution has received a considerable amount of attention, in particular since the development of a source attribution model for salmonellosis by Hald *et al*.[1] The Hald model makes explicit use of *Salmonella* typing data by using differences in the relative frequency of occurrence of *Salmonella* subtypes in individual sources to quantify their contribution to the human disease burden. The Hald model has now successfully been applied to salmonellosis in several EU Member States.[9,10]

We have modified the Hald model to make it more generic and then applied it to New Zealand's major food-borne zoonoses, namely, campylobacteriosis and salmonellosis. Depending on the pathogen, a different subset of modifications was chosen. Those modifications included the inclusion of nonfood sources and a new approach to estimating the prevalence of different pathogen subtypes in a source in the absence of large-scale surveillance.

Campylobacteriosis is the most frequently found food-borne zoonoses in almost all developed countries.[11] However, the development of successful attribution models for this pathogen has been held back by the lack of an appropriate typing scheme as well as its complex epidemiology, which includes environmental reservoirs.[12] With the recent application of multilocus sequence typing (MLST) to *Campylobacter jejuni*, a typing technique has become available, which can be used to investigate the importance of individual components of the food chain as sources of human *C. jejuni* infection.[13] This technique was developed to understand bacterial population structures and is able to show host association for this pathogen.[12,14,15] An additional challenge to source attribution for *Campylobacter* is modeling the role of environmental contamination. Although it is relatively straightforward to estimate the prevalence of a pathogen in a particular type of retail product, such as broiler carcasses, and combine this with the estimated consumption within a population to get an estimate of the exposure through this food source, this approach is hard to apply to the environment as a source of disease. While the environment as a disease source can be thought of as a proxy for unmeasured wildlife sources, it may also be a transmission pathway for pathogens present in livestock sources such as bovines. In order to adapt the Hald model, which originally uses the amount of food source consumed to apportion human cases, changes had to be made to key parameters of the model.

Sparse data were a major problem with our salmonellosis model. The original application of the model makes use of Denmark's extensive surveillance system, and as a consequence point estimates for the prevalence of the different *Salmonella* types in the animal food sources could be used. In New Zealand data were much sparser and an approach was developed to represent uncertainty in the prevalence matrix. In addition, information about prevalences from different studies had to be combined to derive the best possible estimates. In the next section we review the Hald model briefly. We then describe our modifications, which we illustrate by analyzing data on salmonellosis and campylobacteriosis in New Zealand.

In alignment with the Hald model and other approaches, it is assumed that cases with a history of travel in the incubation period have acquired the infection overseas. As a consequence, known travelers were excluded from the attribution process. In our campylobacteriosis model travel status was known for all cases, whereas in the salmonellosis model, where unknown, the status was estimated as proposed by Hald *et al*.

2. THE BAYESIAN RISK ATTRIBUTION MODEL BY HALD *ET AL*.

The Hald salmonellosis model[1] was published in 2004, and it compares the number of human cases caused by different subtypes with their prevalence in different food sources, weighted by the amount of food source consumed. To do this effectively, it requires a heterogeneous distribution of bacterial subtypes among the different animal and food sources. This model is a further development of the so-called Dutch model,[16] a frequentist model, which compares the number of reported human cases caused by a particular bacterial subtype with the relative occurrence of that subtype in each source. By using a Bayesian approach implemented in the WinBUGS software, the Hald model can explicitly include and quantify the uncertainty surrounding each of the parameters. A detailed analysis of the Dutch and the Hald model on Dutch data has been made by Hald *et al*.,[16] and gives comparable results for attribution.

In the model, o_i represents the number of human cases of type i, and the expected number of cases of *Salmonella* type i from source j is denoted by λ_{ij}. Using the parameters defined in Table I, assume that

$$o_i \sim \text{Poisson}(\Sigma_j \lambda_{ij}), \tag{1}$$

Table I. Description of Parameters of Hald Model

Parameter	Description
λ_{ij}	Expected number of cases/year of type i from source j
M_j	Amount of food source j consumed
p_{ij}	Prevalence of type i in source j
q_i	Bacteria-dependent factor for type i
a_j	Source-dependent factor for source j
o_i	Number of human cases of type i

and

$$\lambda_{ij} = M_j p_{ij} q_i a_j. \quad (2)$$

Here, the bacteria-dependent factor q_i combines survivability, virulence, and pathogenicity of the pathogen to estimate the ability of that type to cause disease (measured in cases per dose of bacteria in the population). On the other hand, the source-dependent factor a_j summarizes the ability to act as vehicle for food-borne infections including factors such as the environment provided for the bacteria through storage and preparation (measured in doses of bacteria per kilogram of infected material consumed). Usually, q_i is defined to equal 1 for some reference type, thereby defining implicitly the size of a "dose."

One difficulty of this modeling approach is the disparity between the number of data points and the number of parameters to be estimated. Given I types and J sources, there are $I + J$ parameters (q_i and a_js) but only I independent data points (the observed case totals o_i), so the model is not identifiable: estimates cannot be obtained from the data alone and will be sensitive to the priors used. This nonidentifiability raises the question of whether the output values are more a product of the assumptions made than the data from which they were derived. The pooling of bacterial subtypes into groups with similar characteristics is one way of addressing this problem, as fewer parameters then have to be estimated. This was done in the Hald salmonellosis model by assuming that q_i is of equal value for *S.* Enteritidis and *S.* Typhimurium subtypes. In addition, it was assumed that a_j is equal for some foods, for example, Danish and imported pork. The priors chosen for q_i and a_j

- $q_i \sim$ uniform(0,10)
- $a_j \sim$ uniform(0, 0.01)

were assumed to be noninformative.

Data quality and representativeness of the data have been identified as key determinants of the successful adaptation of this model.[9] However, this has not yet been formally explored and proof of this could be provided by simulating degrees of data quality and representativeness and describing how the model performs in response. In the original model approximately 25% of human cases were assigned to an unknown source. This is a result of the grouping scheme where some *salmonellae* are pooled in a category of "others including not typed." Hald *et al*. concluded that this fraction could be reduced by introducing more subtypes individually into the model.[1]

3. A MODIFIED HALD MODEL

A set of modifications was developed to make the Hald model more generic. This includes a methodology for incorporating uncertainty in the prevalence parameters. We have also taken a different approach to achieving identifiability and to the setting of priors. In addition, approaches were developed to include potentially pathogenic subtypes and to avoid food consumption weights when considering environmental sources. A subset or all of these modifications can be used to apply the model to a particular pathogen and data set. The modifications chosen will depend on the available data quality and the pathogen's epidemiology.

3.1. Modeling Prevalence Uncertainty

In our modified model uncertainty is introduced in the estimates for the prevalence p_{ij} of type i in source j. The original model uses data from Denmark's intensive surveillance system as a justification for keeping source prevalences fixed. In order to make the model more generic and therefore applicable to other countries and pathogens, it was necessary to introduce uncertainty into these estimates. The priors for the p_{ij} were chosen to be independent beta (α_{ij}, β_{ij}) distributions. The parameters α_{ij} and β_{ij} are determined by equating the first two moments of this prior with the first two moments of the posterior distribution obtained from a Bayesian analysis of the prevalence data for each source. Depending on the data, we have developed a standard and a novel approach to estimate prevalence. In the standard approach to the analysis of the prevalence, we assumed that $p_{ij} = \pi_j r_{ij}$, where π_j is the prevalence over all types in source j and r_{ij} is the relative occurrence of type i in the successfully typed isolates from source j. The priors used were $r_{ij} \sim$ Dirichlet (1, $1, \ldots, 1$) and $\pi_j \sim$ beta (1, 1).

Good data on the prevalence of different subtypes in a source for the standard approach may not be available, and we have developed an alternative approach to use data from different studies, including data provided by routine surveillance as well as small-scale surveys. In this novel approach we consider three different types of study: investigating relative prevalence r_{ij}, prevalence over all types π_j, or absolute prevalence $p_{ij} = \pi_j \times r_{ij}$. For ease of exposition we assume only three pathogen types. The three data types are as follows:

1. Typed positives only: observe X_1, X_2, and X_3 samples positive for types 1–3 out of $X_1 + X_2 + X_3 = N_X$.
   The contribution to the likelihood function will be:
$$L_1(r_1, r_2) \propto r_1^{X_1} r_2^{X_2} (1 - r_1 - r_2)^{X_3}. \quad (3)$$
2. Typed positives and negatives: observe Y_1, Y_2, and Y_3 samples positive for types 1–3 out of a total of N_Y samples.
$$L_2(\pi, r_1, r_2) \propto (\pi r_1)^{Y_1} (\pi r_2)^{Y_2} \times (\pi - \pi r_1 - \pi r_2)^{Y_3} (1 - \pi)^{N_Y - Y_1 - Y_2 - Y_3}. \quad (4)$$
3. Untyped positives and negatives (overall prevalence): observe Z positives out of N_Z samples.
$$L_3(\pi) \propto \pi^Z (1 - \pi)^{N_Z - Z}. \quad (5)$$

Putting these all together gives the likelihood function from all available data as:

$$L_1(\pi, r_1, r_2) \propto r_1^{X_1 + Y_1} r_2^{X_2 + Y_2} (1 - r_1 - r_2)^{X_3 + Y_3} \times \pi^{Z + Y_1 + Y_2 + Y_3} (1 - \pi)^{N_Z + N_Y - Z - Y_1 - Y_2 - Y_3}. \quad (6)$$

If we start with independent priors, Dirichlet (1,1,1) for $(r_1, r_2, 1 - r_1 - r_2)$ and beta (1,1) for π, we get independent posteriors (by factorization theorem):

$$(r_1, r_2, 1 - r_1 - r_2) \sim \text{Dirichlet} \times (X_1 + Y_1 + 1, X_2 + Y_2 + 1, X_3 + Y_3 + 1), \quad (7)$$

$$\pi \sim \text{beta}(Z + Y_1 + Y_2 + Y_3 + 1, N_Z + N_Y - Z - Y_1 - Y_2 - Y_3 + 1) \quad (8)$$

Where data are available in different time periods, for example, in consecutive years, it may be preferable to smooth the observed numbers of cases over time using a weighted moving average. These posteriors incorporating the available prevalence information could now be used as priors in the source attribution program. We have found the full scheme to be too complex for our WinBUGS updater, so we have implemented it as a two-stage process as follows.

First Equations (7) and (8) were implemented in WinBUGS and posterior means and standard deviations (*SD*) for the prevalences p_{ij} were obtained. The parameters of a beta distribution (α_{ij} and β_{ij}) were chosen to match the mean and *SD* of each prevalence. In order to avoid convergence problems for very small values of α_{ij}, we enforced a minimum $a = 1$ and set the corresponding β_{ij} to match the mean only. The values of α_{ij} and β_{ij} for the different prevalences p_{ij} were then used to specify beta priors in the risk attribution model for each source.

3.2. Splitting Data into Different Time Periods

An alternative, or complementary, approach for achieving identifiability is to divide the observation period into a number of intervals and to estimate prevalences separately in each time interval t. Equations (1) and (2) now become:

$$o_{it} \sim \text{Poisson}(\Sigma_j \lambda_{ijt}),$$
$$\lambda_{ijt} = M_{jt} p_{ijt} q_i a_j,$$

where all data values and parameters are assumed to change between time periods, except for q_i and a_j, which are assumed to be constant over time. This increases the number of independent data values while keeping the number of parameters to be estimated constant, and will lead to identifiability provided the prevalences p_{ijt} do vary with time.

3.3. Hierarchical Model for Bacterial Parameters

To achieve identifiability in the model, the number of q_i parameters must be reduced to be less than the number of observations o_i. Hald *et al.* do this by assuming that some of the q_i are equal. We prefer to model the q_i hierarchically as random observations from a hypothetical distribution of bacterial characteristics. We used a lognormal distribution $\log(q_i) \sim N(0, \tau)$, where τ is a parameter controlling the variation in characteristics between types. A prior distribution is needed for τ; we used a fairly diffuse gamma prior $\tau \sim \text{gamma}(0.01, 0.01)$. We conducted a sensitivity analysis and found that variations in the choice of prior did not significantly influence the results.

The mean of 0 on the log scale parallels Hald *et al.*'s approach of fixing one q_i to equal 1. This random effects approach allows for the possibility

that some parameters may be similar, as in the Hald model, but without forcing them to be equal. If genuine prior information is available on relative characteristics, this can be incorporated. The model is now identifiable because the I virulence parameters are replaced by a single τ parameter describing the virulence distribution. This is particularly useful if there are a large number of subtypes.

3.4. Exponential Prior for Source-Specific Parameters

We also changed the priors for the a_j parameters. Independent prior distributions for the source-dependent factors were chosen as $a_j \sim$ exponential (λ). This prior is fairly uninformative, but prevents the a_j from becoming too large, while not specifying a strict upper boundary as in the Hald model. An alternative would be to induce priors on the a_j by speculating on the likely observed number of cases for an "average" pathogen type given a fixed prevalence and amount consumed of source j. The priors for a_j will depend on the inclusion of food consumption weights and their unit of measurement. In our analysis we investigated the sensitivity of the results to the value of λ chosen for the prior.

3.5. Avoid Food Consumption Weights

For some pathogens such as *C. jejuni* the environment is a well-known source of infection[12] and adjustment has to be made to include it into the model, as the original approach was focused on animal food sources. The Hald model uses a food-consumption-weighting factor M_j in its specification. To include the environment in the model, we removed M_j because weighting exposure to the environment in a comparable way to consumption of a food source (as measured in, e.g., kilogram of the product produced for consumption) could not be achieved in a sensible way and the M_j is not essential to the model. With the M_j in the model the related food-source-dependent factor a_j can be interpreted as the relative ability of a certain food source to cause disease. By removing M_j the information about the amount of food consumed in a population is not explicitly included in the model but will be absorbed in the values for a_j. In general, the estimated values for a_j (and q_i) are simply multiplication factors to arrive at the most probable solution given the observed data and their size gives an idea about the different ability of the food sources to transmit disease (or the bacteria subtypes to cause disease). In the modified model a high a_j, for example, may reflect a high exposure (e.g., a large market share), but not necessarily a high ability of the individual food source to cause disease, as a consequence the a_j are less meaningful and less comparable. This modification is not unique to our models and has also been applied to the adaptation of the Hald model to salmonellosis in Sweden.[10] We extend this work by outlining the theory behind this adaptation and describe the consequences for the interpretation of the results.

3.6. Including Potentially Pathogenic Subtypes

The Danish data used in the Hald model consist of only one particular type of data, *Salmonella* types that have been found in both human cases and at least one food source. The Danish data have a pooled group "other *Salmonella* including not typable isolates" as well as the major individual *Salmonella* types. As a consequence, the origin of approximately 25% of the human cases was classified as "unknown."[1] However depending on the resolution of the typing method, a data set can consist of three data types: subtypes (STs) occurring in both human cases and at least one of the sources (like in the Danish data set), those only occurring in humans, and those only occurring in at least one of the sources but not in humans. We chose to also include the third type, the potentially pathogenic STs into the model. These can be found in the sources but have not (yet) been detected in humans. We assume that these subtypes are potentially pathogenic but rarely occur in humans. We did not include human types that have not been found in any of the sources. In the absence of information relating them to any of the major sources we are assuming, those types may have come from unknown, possibly exotic, sources. In consequence, the model only attributes cases of subtypes that have been detected in a source. To attribute types undetected in a source, inferences from genotypic relatedness may be used as proposed by Wilson *et al.*[17] However, this is only applicable where genotyping data are being used, and therefore has not been applied as yet to *Salmonella*, which is routinely investigated by phage typing.

4. APPLICATION OF THE MODIFIED HALD MODEL

4.1. Campylobacteriosis

4.1.1. Details of Campylobacteriosis Model

Data for campylobacteriosis source attribution were generated by a sentinel surveillance study for *C. jejuni* in the Manawatu region of New Zealand

conducted between 2005 and 2008[18] and the allocation is based on 481 MLST typed human cases. The Manawatu study consists of structured parallel studies of isolates from domestic human cases of *C. jejuni* and environmental and food sources in a defined geographical area over a three-year period. Retail chicken samples were used to represent poultry sources since in New Zealand chicken represents 95% of poultry meat consumed. Details of the study are discussed by French *et al*.[18] and further work is underway to differentiate between individual poultry sources.

All isolates in the data set were completely typed and each *C. jejuni* subtype was modeled individually.

The following modifications were applied to this pathogen:

- Modeling prevalence uncertainty using the standard approach
- Using a hierarchical model for bacterial parameters
- Using an exponential prior for source-specific parameters
- Avoiding food consumption weights
- Including potentially pathogenic subtypes

Samples from the posterior distribution for this model were obtained using Markov chain Monte Carlo (MCMC) techniques, run in the software WinBUGS 1.4 called from R 2.5.1 (using the R2WinBUGS package). The WinBUGS code was developed from the original code by Hald *et al*. A total of 10,000 samples were taken from five independent Markov chains with widely dispersed starting values, after a burn-in period of 2,000 iterations for each chain with a thinning of 10. The prior distribution for τ was assumed to be $\tau \sim$ gamma (0.01, 0.01).

The sensitivity of the model for two different priors for the food source factor a_j, the standard prior $a_j \sim$ exponential (0.002) and a less restricting prior $a_j \sim$ exponential (0.01), was assessed. The model's sensitivity against different groupings of the sources was tested. The source "poultry" was split into the major three suppliers and the source "ovine" sample was split into "retail" and "on-farm samples." In a second step we assessed the model's sensitivity to the inclusion of potentially pathogenic *C. jejuni* subtypes.

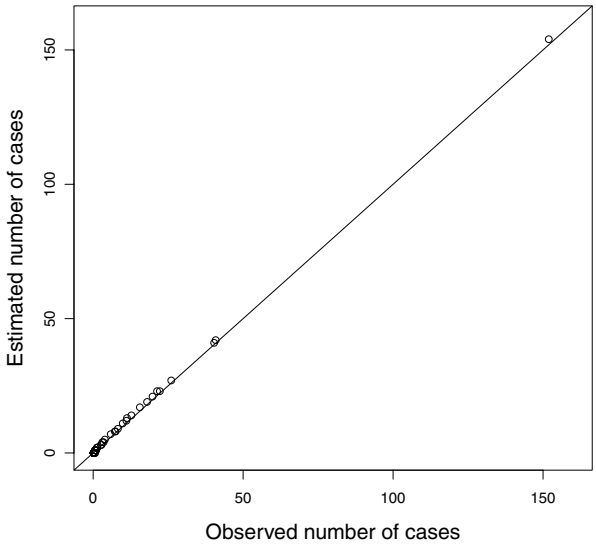

Fig. 1. Plot of the observed (o_i) and expected cases (λ_i) for the individual *Campylobacter* subtypes.

4.1.2. Results of Campylobacteriosis Model

The risk model for *C. jejuni* attributes 474 human cases, a close match to the observed number of 481 cases. Overall observed (o_i) and expected cases (λ_{ij}) for the individual *Campylobacter* subtypes were in good agreement (Fig. 1). Out of the 474 cases, 379 were attributed to poultry (80%), 48 cases to bovine (10%), 44 cases to ovine (9%), and 4 cases to the environment (1%) (Fig. 2A). Including uncertainty in the prevalence estimates widens the credible intervals of the attribution estimates and mainly affects the estimates for bovine and poultry. The bovine estimate is affected the most; its point estimate changes from 48 to 68 cases and the upper credible interval limit increases from 160 to 180 (Fig. 2A). The effect of a change of prior from $a_j \sim$ exponential (0.002) to $a_j \sim$ exponential (0.01) was minor; predictions were within 15% of the standard approach (Fig. 2B).

To test the sensitivity of the model to changes in the underlying assumptions several alternative models were developed (Table II). The impact of the choice of model on the source attribution estimates is documented in Fig. 3. The model showed a low sensitivity to changes in the source grouping. It was able to distinguish between ovine on-farm and retail exposure without affecting the overall estimate and similarly could distinguish between the three major New Zealand poultry suppliers.

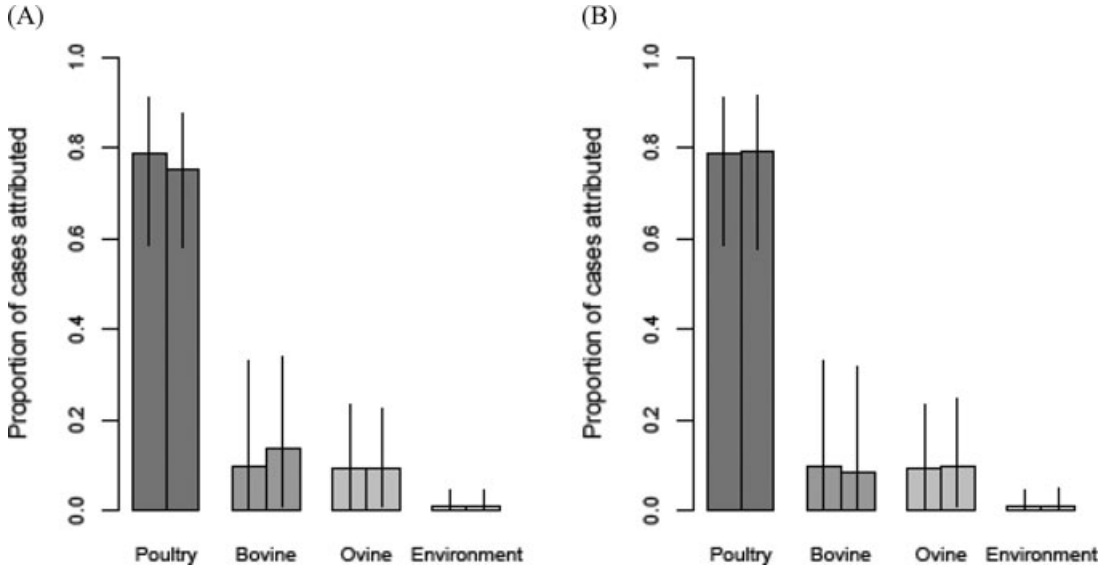

Fig. 2. Attribution results based on 481 campylobacteriosis cases in the Manawatu region of New Zealand to several sources. Median proportion of cases attributed to each source with 95% Bayesian credible intervals. (A) The graph shows results from the modified Hald model with (on the left) and without uncertainty in the prevalence matrix (on the right). (B) The graph shows the sensitivity of the modified Hald model for different priors for a_j. The standard model using $a_j \sim$ exponential (0.002) is represented in the left column and $a_j \sim$ exponential (0.01) in the right column.

Table II. Description of Standard and Alternative Models to Test Robustness of Campylobacteriosis Model

Model Number	Model Name	Source Grouping	Inclusion of Potentially Pathogenic Subtypes
I	Standard	Poultry Bovine Ovine Environment	Yes
II	No extra types	Poultry Bovine Ovine Environment	No
III	Split ovine	Poultry Bovine Retail ovine On-farm ovine Environment	Yes
IV	Split chicken	Poultry supplier 1 Poultry supplier 2 Poultry supplier 3 Bovine Ovine Environment	Yes

It was, however, more sensitive to the inclusion of potentially pathogenic *C. jejuni* subtypes, which affected all estimates, except the one for ovine sources.

4.2. Salmonellosis

4.2.1. Details of Salmonellosis Model

In the New Zealand dataset, the *Salmonella* subtype was known for all typed human cases; however, the resolution of subtypes in the food source was lower. As a consequence, human subtypes were reported in the same categories as the types from the food sources. This was based on the scheme commonly used in New Zealand for surveillance purposes, which includes a total of 16 types, one of which is a pooled category of less frequently occurring subtypes. This group, called "other *Salmonella*," was very diverse, and initially included a total of 95 different subtypes, many of which were associated with foreign travel and have not been detected in any source in New Zealand. Estimates of food source prevalence were calculated for 13 *Salmonella* types and together with the occurrence of these 13 types in human cases used to attribute cases. Three of the subtypes have not been detected in any food sources and were therefore excluded from the model, based on the conclusions drawn in Section 3.6. These cases account for 2–4% of all domestic, sporadic cases annually.

Human case data were derived from New Zealand's national surveillance system EpiSurv.[19] For the food sources in the absence of large-scale

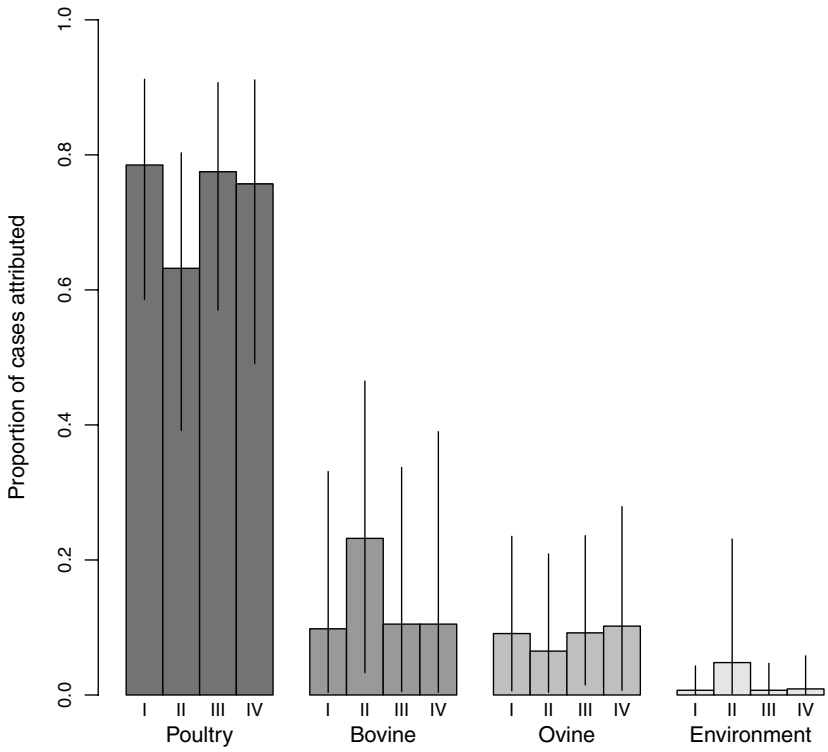

Fig. 3. Sensitivity analysis of source attribution results to different grouping of sources and inclusion of minor *Campylobacter* types. The different models are described in detail in Table II. Models I to IV are shown from the left to the right and summary estimates are shown where sources were split up. Median number of cases attributed to each source with 95% Bayesian credible intervals.

high-quality surveillance data information from different studies was used to estimate the prevalences of the different *Salmonella* types in the food sources. These included data provided by routine surveillance as well as small-scale surveys. The amount of data was similar for all sources except pork, for which data were very sparse and in parts unrepresentative. Pork is the only meat product in New Zealand that is not routinely tested and only very few isolates are typed from pork annually. The consequences of this are being considered in the discussion of our results.

As an example for the novel prevalence estimation approach, we are presenting here results for *Salmonella* in beef and veal in 2003. For this source three different data sets were available:

1. Data summarizing isolates submitted to the National Reference Laboratory for *Salmonella* (X),
2. Data from a major retail survey conducted in 2003 and 2004 (Y_1),
3. Data from the National Microbiological Database (NMD), which standardizes New Zealand's official export assurances to overseas markets (Y_2).

These three data sets were combined into a data matrix (Table III) and consecutively analyzed as described in Section 3.1.

The following modifications to the Hald model were applied for this pathogen:

- Modeling prevalence uncertainty using the novel approach
- Splitting data into different time periods
- Using an exponential prior for source-specific parameters
- Using a hierarchical model for bacterial parameters

We modeled a combined estimate for the years 2002–2004 for salmonellosis, using individual year data on human cases and source prevalence. A pooled (three-year) value for each q_i and a_j was estimated, while source attribution estimates (λ_j) were calculated individually for each year. As in the Hald model, equality in the type-specific bacteria-dependent parameters for *S.* Typhimurium subtypes was assumed.

The model was fitted using MCMC techniques with the software WinBUGS 1.4.3. The code was developed from the original code by Hald *et al*. Five independent Markov chains, with widely dispersed starting values, were run for 80,000 iterations after

Table III. Data Matrix for *Salmonella* Subtype Prevalence Estimation in Beef and Veal in 2003

	Salmonella Subtype *i*														
	1	2	3	4	5	6	7	8	9	10	11	12	13	Total Isolates	Total Samples
X (2002)	53	3	0	5	24	0	0	33	13	27	15	11	61	245	–
X (2003)	30	2	3	1	10	5	0	26	17	0	25	4	78	201	–
X (2004)	0	0	0	1	0	0	0	1	0	0	0	0	0	2	–
X[a]	**33.1**	**2.1**	**2.3**	**1.8**	**12.3**	**3.8**	**0**	**26.2**	**15.4**	**5.4**	**21.8**	**5.2**	**70.7**	**200.1 (Nx)**	**–**
Y_1	0	0	0	1	0	0	0	1	0	0	0	0	0	2	294
Y_2 (2002)	6	0	0	0	1	0	0	0	0	1	1	1	2	12	2,138
Y_2 (2003)	4	0	0	0	4	0	0	1	3	0	0	2	2	16	1,768
Y_2 (2004)	6	0	0	0	2	0	0	0	0	0	1	0	2	11	1,549
Y_2[b]	5	0	0	0	2.8	0	0	0.5	1.5	0.3	0.5	1.3	2	13.8	1,805.8
Y[c]	**5**	**0**	**0**	**1**	**2.8**	**0**	**0**	**1.5**	**1.5**	**0.3**	**0.5**	**1.3**	**2**	15.8	**2,099.8 (Ny)**

[a]X is calculated by combining data from 3 years using weighted moving averages. These are typed isolates only.
[b]Y_2 is calculated by combining data from 3 years using weighted moving averages. These are typed isolates with prevalence data.
[c]Y is calculated by combining Y_1 and Y_2.
Note: Included data X, Y_1, and Y_2, and methods are discussed in text.

a burn-in period of 40,000 iterations, with a thinning of 50. Convergence was monitored using the method developed by Gelman and Rubins.[20] The length of the chain was determined by running sufficient iterations to ensure that the Monte Carlo errors for each parameter were less than 5% of the posterior standard deviation.[20] The prior distribution for τ was assumed to be $\tau \sim$ gamma (0.01, 0.01). As for the campylobacteriosis model the sensitivity of the model for two different priors for a_j, namely, $a_j \sim$ exponential (0.002) and $a_j \sim$ exponential (0.01), was assessed.

4.2.2. Results of Salmonellosis Model

Table IV shows median prevalence estimates and corresponding Bayesian credible intervals for *Salmonella* (p_j) and individual *Salmonella* subtypes (p_{ij}) in beef and veal in 2003 as estimated by our novel approach.

Based on a total of 963 observed cases, the risk model apportions an estimated 981 human cases in 2003. As in our campylobacteriosis model, observed (o_i) and expected cases (λ_{ij}) for the individual subtypes were very close in value (data not shown). The majority of cases were attributed to pork (60%, 591 cases) followed by poultry (21.2%, 209 cases) and beef and veal (11.5%, 113 cases). Eggs and lamb and mutton are estimated to be minor sources of infection with 3.2% and 1.4% of cases apportioned (Fig. 4). The sensitivity of the model against changes in the prior for a_j was tested. Changing the prior caused a minor change in the estimates for the two minor sources eggs, poultry, and lamb and mutton (Fig. 4).

Table IV. Prevalence Estimates in Percent for *Salmonella* in Beef and Veal in 2003

	Median	2.5% CI[a]	97.5% CI[a]
p_i for *Salmonella* subtype *i*			
S. Brandenburg	0.133	0.073	0.227
S. Enteritidis PT 9a	0.009	0.002	0.028
S. Heidelberg	0.01	0.002	0.029
S. Infantis	0.012	0.003	0.033
S. Other/Unknown	0.054	0.026	0.103
S. Saintpaul	0.015	0.005	0.038
S. Thompson	0.002	0	0.014
S. Typhimurium DT 1	0.097	0.052	0.172
S. Typhimurium DT 101	0.06	0.03	0.112
S. Typhimurium DT 135	0.022	0.008	0.049
S. Typhimurium DT 156	0.079	0.041	0.142
S. Typhimurium DT 160	0.024	0.009	0.054
S. Typhimurium other/unknown	0.253	0.145	0.41
Overall prevalence			
p	0.7813	0.4591	1.221

[a]Bayesian credible interval.

5. DISCUSSION

5.1. Improving Identifiability

The full model as specified by Hald *et al*. is overparameterized. This was solved in the original model by assuming equality in some of the type-specific bacteria-dependent and food-source-specific parameters. Our solution involves using a hierarchical model for the bacterial parameters and combining

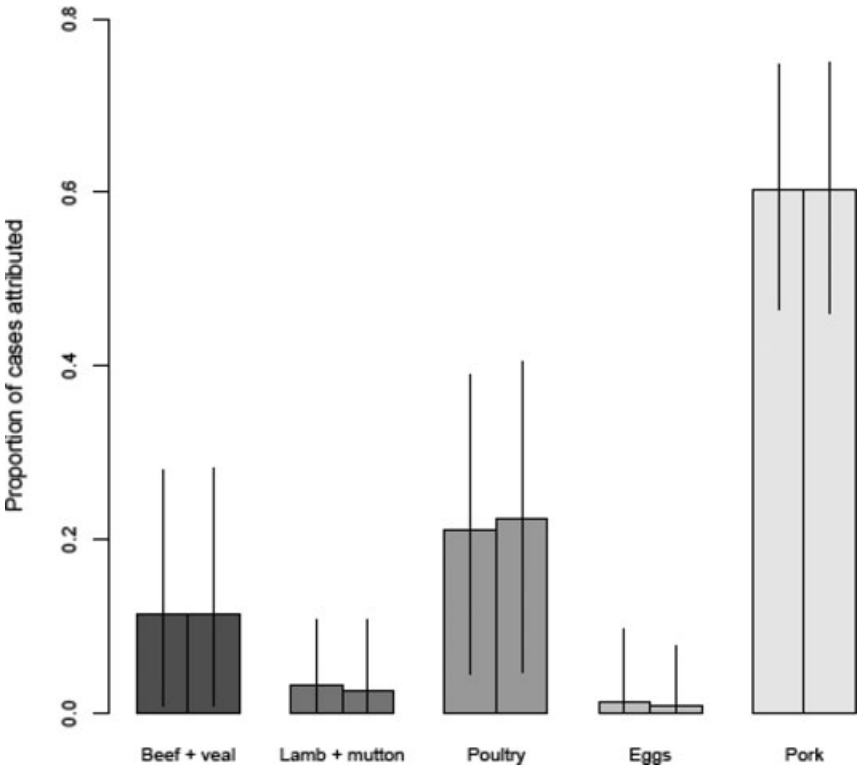

Fig. 4. Attribution of 891 human salmonellosis cases in New Zealand in 2003. Median proportion of cases attributed to each source with 95% Bayesian credible intervals. The graph shows the sensitivity of the modified Hald model for different priors for a_j. The standard model using $a_j \sim$ exponential (0.002) is represented in the left column and $a_j \sim$ exponential (0.01) in the right column.

data from different time periods to achieve identifiability.

It is biologically plausible to assume that the q_i for all pathogen subtypes come from a common distribution. A major advantage of this approach is that no assumptions about equal value for any of the q_is have to be made. In our salmonellosis approach it was possible to model the q_i hierarchically while also assuming equality for *S.* Typhimurium and *enteritidis* subtypes, as in the original approach. This could not be done for campylobacteriosis, as to date there is no evidence that would justify a grouping of MLST types. It was assumed in the Hald salmonellosis model that the source-dependent factor a_j is equal for some foods, for example, for Danish and imported pork. We have not been able to identify sources for which this assumption could be made in any of our models.

To further address identifiability the modified model, as applied to salmonellosis in New Zealand, used human case and animal food source data from individual years while estimating a pooled q_i and a_j over a three-year period. This change improved the ratio of data points and parameters and considerably improved the performance of our model. The assumption behind this is that those factors are constant in time. This would, as an example, include the ability of beef to cause disease, which results from a variety of factors such as the survivability of this pathogen in this food source. The assumption that a_j and q_i are constant in time may not always be accurate, in particular for the q_i, which has to be carefully considered depending on the context. Prevalence of endemic *Salmonella* types is often observed to fluctuate in time and this pathogen is also well known for emerging types, causing outbreaks within populations.[21] Changes in consumer behavior such as an increased consumption of raw eggs would have an impact on the true value of a_j. By careful consideration, the above modeling of pooled estimates can be a way of improving model performance.

5.2. Choice of Prior for q_i and a_j

The original Hald model uses universal uniform priors for q_i and a_j, which contain very little information about the parameters. In our modified approach q_i is modeled as a random effect and its variation is controlled by a hyperparameter τ. By taking this approach, we are assuming a common distribution for all q_i, which adds to our knowledge and improves the model. A valuable extension to this, particularly

suitable for salmonellosis, would be to take a nested hierarchical approach, allowing, for example, for random effects between *S.* Typhimurium and *S.* Enteritidis subtypes. The choice of prior for τ may be influential, so a sensitivity analysis should be carried out to investigate the effect of different priors.

We used an exponential distribution to model a_j. Our sensitivity analysis for both models showed a low sensitivity of the estimates to the exact exponential distribution chosen. A major advantage of taking a Bayesian approach is that by running the model for different periods, the priors for q_i and a_j could be updated and improved by including estimates of a previous model as prior for the updated model. At the moment, the model does not address the possibility of interaction between the food a_j and the bacteria-related factor q_i. This would allow for the biologically plausible possibility that certain subtypes are more or less likely to survive and cause disease, dependent on the food source they appear in. There is evidence in the literature for such an interaction.[22,23]

5.3. Introduction of Uncertainty in Prevalence Matrix

In general, for a full Bayesian approach, uncertainty in the prevalence p_{ij} estimates should be included. In particular, in the absence of intensive surveillance data as used in the original Hald model, uncertainty around the estimates cannot be ignored as our results would otherwise overestimate the level of precision. Technically, the inclusion of uncertainty around the prevalence estimates in the model removes zeros from the prevalence matrix, possibly increasing the number of parameters and making the model more difficult to fit. However, if assuming zero prevalence is incorrect then the model will be misspecified. We have evaluated the effect of introducing uncertainty in the prevalence p_{ij} estimates for our campylobacteriosis model. As a consequence, in this model attribution confidence intervals were moderately wider and point estimates slightly changed for two of the sources. By incorporating this additional layer of uncertainty, the model now more closely reflects the true uncertainty in the risk estimates.

5.4. Novel Approach to Prevalence Estimation

In general, the Hald model requires intensive monitoring of all relevant sources and human cases followed by the application of discriminatory epidemiological typing methods.[16] The general concept is, however, applicable where the consequences of using biased or sparse data are being considered. First, a distinction has to be made between data quality (or representativeness) and data quantity (or sparsity). Sparse data will result in a low precision of the prevalence estimate and as a consequence a higher uncertainty in the whole model. On the other hand, unrepresentative data will lead to bias in the estimates and this effect will need to be carefully discussed. New Zealand's *Salmonella* animal food source data set has been created from several fragmented data sources from different origins. These data can clearly never be as representative as large-scale standardized surveillance. In practice, however, even if a known major source of a disease is closely monitored, this is almost never the case for all known sources, in particular for minor sources such as game birds. The approach we developed has enabled us to formally estimate prevalence in a source, including the uncertainty surrounding our estimate. By objectively combining several sets of existing information, we could maximize our knowledge and come up with the best possible estimate. Applying this approach has also helped us to identify data gaps and to evaluate existing surveillance systems. A way forward would be to address data quality and quantity in the analysis, by incorporating them in the prior for a_j and p_{ij}.

5.5. Splitting Data into Different Time Periods

To achieve identifiability, the observation period can be divided in a number of intervals and prevalence estimated separately while q_i and a_j are assumed to be constant over time, as we have illustrated in our salmonellosis model. In our campylobacteriosis model we estimate the contribution of different sources to the human disease burden over a three-year period, while the original Hald model produces an annual estimate. Overall, the model offers the opportunity for dynamic attribution modeling in time and this is planned to be further explored by our research group. In particular for disease with a seasonal component, such as campylobacteriosis, this would be of great interest and could greatly contribute to our understanding of disease dynamics.

5.6. Avoid Food Consumption Weights and Modeling Environmental Sources

We have reasoned that it is without major consequences to avoid food consumption weights by taking the M_j out of the model, but it will make the a_j less comparable. Removing the M_j may be

necessary when no reasonable and comparable consumption estimates are available, for example, when environmental sources are considered. It is worth mentioning that in the original Hald model M_j is either measured in tons (for meat sources) or counts (for eggs).

Since disease transmission from a farm animal source may be via both food and nonfood pathways, farm animal sources may be better represented taking this approach. Estimating the contribution from farm animal reservoirs (or amplifying hosts) by using the distribution of genotypes present in both food and on-farm fecal material, we attempt to capture the contribution from both food and nonfood pathways. The importance of such nonfood pathways is underlined by assessments of exposure to *Campylobacter*.[24] This adaptation offers many opportunities to model nonfood sources and has been applied previously to include travelers as a disease source.[10]

Including specific environmental pathways as additional "sources" in the model requires careful consideration. In the case of our wild bird environmental source, this represents a nonfood pathway in which we assume humans are directly exposed to fecal material from a reservoir of nonfood-producing wild animals. Our water environmental source is, however, not strictly speaking a reservoir or amplifying host itself, but represents an exposure pathway that is nonfood, but could contain fecal material from both food-producing and nonfood-producing animals. Therefore, some of the human cases attributed to the separate environmental category could be considered as additional cases arising from farm animals via a nonfood pathway. However, although some ruminant-associated and poultry-associated genotypes were isolated from environmental water, the majority belonged to genotypes associated with wildlife. We therefore conclude that, although some of the cases attributed to environmental water are likely to originate from nonwildlife sources, this is likely to be a very small fraction. The most common *C. jejuni* subtype was ST-2381—this genotype has not been identified anywhere other than in New Zealand water.[18] In addition, wild birds and water isolates showed the lowest similarity with human isolates.[18] Given that only 1% of cases are attributed to environmental sources, the effect of how environmental pathways and reservoirs are classified and treated in the model is likely to be negligible. Our approach has included such sources; however, this needs to be better understood and more work is underway to improve our knowledge of the epidemiology of *Campylobacter* in wildlife.

5.7. Including Potentially Pathogenic Subtypes

We chose to include potentially pathogenic STs into the model that occur in the sources but have not (yet) been detected in a human sample. The model assigns each of these potentially pathogenic subtypes a low probability of causing a case. This approach assumes that these types are not apathogenic but rare human pathogens. By introducing these potentially apathogenic types into the model, we account for the proportion of subtypes from a source that are proven and common human pathogens. This further advances our knowledge of the overall importance of a source and this is reflected in the estimate. However, if these types are truly apathogenic, the few estimated cases for these subtypes will be an overestimate. Due to the low number of attributed cases this accounts for, this will not affect the model much and is outweighed by the insight we gain into the relative frequency of common pathogens in the sources. As an example, only 26% of the *C. jejuni* isolates typed from environmental water were types that were also found in our human samples, compared with 92% of isolates from poultry.[18] In our campylobacteriosis model we evaluated the sensitivity of the model using this approach. A moderate change in the estimates for some sources could be observed. A moderate increase of the cases attributed to sources with a high proportion of types detected in humans (poultry in our example) and a reduced number of cases attributed to a source with a low proportion of types detected in humans (water in our example) could be observed.

5.8. Campylobacteriosis Model

We have, for the first time, extended the Hald model to a pathogen other than *Salmonella*. Our campylobacteriosis model produces outputs similar to those of other approaches[25,26] and further work is on the way to use the results from these different models as cross-validation of our results. Our modified model shows low sensitivity against changes in the prior for a_j, as well as different grouping of the sources, which suggests a stable model. The application of the model has been supported by the quality of the available data set as well as application of a new *Campylobacter* typing technique, which provides information on host association. Poultry has

been identified as the major source of human campylobacteriosis due to *C. jejuni* in New Zealand, causing an estimated 80% of human cases.

The importance of bovine and ovine sources in disease transmission has been highlighted with an estimated 10% and 9%, respectively, of human cases attributed to these sources. Given the relatively low prevalence and pathogen counts in ruminant meat samples[18,27] and the age and spatial distribution of ruminant-associated human cases,[18] they are more likely to be the result of environmental and occupational, rather than food-borne, exposures, which will require tailored control strategies. However, our model alone cannot readily distinguish between food and nonfood pathways arising from farm animal reservoirs. Further inference can only be gained from additional epidemiological and microbiological data. This underlines the importance of a holistic approach to disease prevention when multiple transmission pathways exist for an individual source. A combination of environmental, occupational, and food safety interventions will be necessary to reduce the number of cases from bovine and ovine sources.

Our risk attribution results do have wide credible intervals, which is only in part caused by the uncertainty in our prevalence matrix. First, these wide confidence intervals may reflect *C. jejuni*'s complex epidemiology and the resulting uncertainty about disease origin. Second, our allocation was based on 481 human cases compared with 3,268 cases in the Hald salmonellosis model. An increase of human cases may be a way forward to increase precision of the estimates.

The model is focused on the main human pathogen *C. jejuni*, which is responsible for an estimated 90% of human campylobacteriosis cases and differs in its epidemiology and infection pathways from other *Campylobacter* species.[28] Studies including the rarer species of *Campylobacter* will be needed to model the total burden of human campylobacteriosis in New Zealand.

5.9. Salmonellosis Model

We have adapted the Hald salmonellosis model to New Zealand. The major challenge we encountered was the data available to estimate prevalence of *Salmonella* subtypes in the food sources. In any of the three years the risk model has identified pork as the major source of human salmonellosis followed by poultry and beef and veal. This contradicts the fact that it is commonly believed that pork consumption is not an important animal food pathway for this disease in New Zealand due to the low prevalence and rare isolation from domestic pork and live pigs.[29] However, in the European Union pork is often implicated as a source and reservoir of human salmonellosis and commonly involved in outbreaks.[30] Given that the data for pork were much sparser and more biased than for any of the other sources, the results have to be interpreted with care as this may have had a major influence on the number of cases attributed. If only few isolates are typed from a source and these include the major human types, the sparse data will lead to an overestimate of the importance of that source. Not including a major source in the model may, however, cause severe bias due to cases from pork being wrongly assigned to other sources. It is therefore preferable to include the source into the model, fill existing data gaps, and take the poor data quality into account when interpreting the results from the model. Layer hens (eggs) have been identified as a major source of the disease elsewhere,[1] but due to the absence of particular pathogenic *Salmonellae* in New Zealand eggs this is likely not to be the case in New Zealand and very few cases appear to be attributable to eggs. We believe that the ranking of poultry, beef and veal, lamb and mutton, and eggs by our model has provided us with valuable information about the contribution of these sources to the disease burden in New Zealand. Attribution estimates for salmonellosis in New Zealand include wide credible intervals. We believe this is a result of the large remaining uncertainty in the prevalence matrix and estimates will be more precise once more data become available.

Due to the differences in the surveillance scheme between New Zealand and Denmark, we have been able to allocate all cases to a food source. A group exists in our data set, which pools minor, but completely typed, *Salmonellae* subtypes together. In the Danish data the pooled group also contains nontyped isolates and as a consequence the source of these types is unknown. We make the assumption that the minor types, which in combination account for 8–11% of domestic, sporadic human cases annually, can be proportionally attributed between the sources. If data become available, it is preferable to assign as many individual subtypes as possible.

The original Hald model did not converge with our salmonellosis data set. This has also been observed in other countries.[9] After modifying the model convergence was achieved, providing estimates of source attribution. The validity of these estimates is highly dependent on the quality of the data used, and sources of bias should be taken into

consideration when interpreting the results. This is illustrated by the apparent high attribution of pork in our study, which is likely to be biased by poor-quality data on prevalence for this source. Therefore, a logical next step for source attribution of salmonellosis in New Zealand will be improving on the quality of the data in these food sources. Although we have concerns about the validity of the source attribution estimates, the application of the modified Hald model to salmonellosis was a useful exercise to better understand the model, to identify and address obstacles to its adaptation, and has led the way to its extension to another pathogen.

6. CONCLUSIONS

We have extended the Hald model for a more generic application by creating a set of modifications that can be individually applied to the model depending on the context. These modifications improve the model's identifiability by various means such as by using a hierarchical model for the bacterial parameter q_i. We have modified the model to consider the contribution from environmental, nonfood pathways, including putative wildlife sources, and applied this modification to campylobacteriosis in New Zealand. In addition, novel methods have been developed to introduce uncertainty around the prevalence estimates and to thereby enable the application of this approach in the absence of a very intensive national surveillance system for all major sources of disease.

ACKNOWLEDGMENTS

We thank all members of the Hopkirk Molecular Epidemiology Team (in particular A. C. Midwinter and J. M. Collins-Emerson), Environmental Science and Research (ESR, in particular P. Carter and C. Kliem), MidCentral Health, Public Health Services, MedLab Central, and the New Zealand Food Safety Authority (NZFSA, in particular the Science Group and S. Wagener) for their contributions. We thank T. Hald and S. M. Pires for providing the code for the Hald model and for their support. In addition, we would like to gratefully acknowledge Professor M. Hazelton's (Massey University) contribution to our discussions and Jonathan Marshall's (Hopkirk Institute) help with the graphs. Financial support was provided through the New Zealand Food Safety Authority projects FDI/236/2005 (enhancing surveillance of potentially food-borne enteric diseases in New Zealand: human campylobacteriosis in the Manawatu) and FSPD/83/2004 (Ph.D. human campylobacteriosis).

REFERENCES

1. Hald T, Vose D, Wegener HC, Koupeev T. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. Risk Analysis, 2004; 24:255–269.
2. Batz MB, Doyle MP, Morris JG, Painter J, Singh R, Tauxe RV, Taylor MR, Wong D. Attributing illness to food. Emerging Infectious Diseases, 2005; 11:993–999.
3. Havelaar A, Braeunig J, Christiansen K, Cornu M, Hald T, Mangen MJJ, Molbak K, Pielaat A, Snary E, Van Pelt W, Velthuis A, Wahlstroem H. Towards an integrated approach on supporting microbiological food safety decisions. Zoonosis and Public Health, 2007; 103–117.
4. Doyle MP, Erickson MC. Emerging microbiological food safety issues related to meat. Meat Science, 2006; 74:98–112.
5. Adak GK, Cowden JM, Nicholas S, Evan HS. The Public-Health Laboratory Service national case-control study of primary indigenous sporadic cases of *Campylobacter* infection. Epidemiology and Infection, 1995; 115:15–22.
6. Eberhart-Phillips J, Walker N, Garrett N, Bell D, Sinclair D, Rainger W, Bates M. Campylobacteriosis in New Zealand: Results of a case-control study. Journal of Epidemiology and Community Health, 1997; 51:686–691.
7. Gillespie IA. Milkborne general outbreaks of infectious intestinal disease, England and Wales, 1992–2000. Epidemiology and Infection, 2003; 130:461–468.
8. Gillespie IA, O'Brien SJ, Frost JA, Adak GK, Horby P, Svan AV, Painter MJ, Neal KR; Campylobacter Sentinel Surveillance Scheme Collaborators. A case-case comparison of *Campylobacter coli* and *Campylobacter jejuni* infection: A tool for generating hypotheses. Epidemiology and Infection, 2002; 128:111–118.
9. Pires SM, Nichols G, Whalstrom H, Kaesbohrer A, David J, Spitznagel H, Van Pelt W, Baumann A, Hald T. *Salmonella* Source Attribution in Different European Countries. Food Micro, Aberdeen, 2008.
10. Wahlstroem Anderson H, Plym-Forsell Y, Pires LSM. Source Attribution of Reported Human Salmonella Cases in Sweden. Food Micro, Aberdeen, 2008.
11. Oberhelman RA, Taylor DN. *Campylobacter* infections in developing countries, Nachamkin I, Blaser MJ (eds). Campylobacter. Washington: American Society for Microbiology, 2000.
12. French N, Barrigas M, Brown P, Ribiero P, Williams N, Leatherborrow H, Birtles R, Bolton E, Fearnhead P, Fox A. Spatial epidemiology and natural population structure of *Campylobacter jejuni* colonizing a farmland ecosystem. Environmental Microbiology, 2005; 7:1116–1126.
13. Dingle KE, Colles FM, Ure R, Wagenaar JA, Duim B, Bolton FJ, Fox AJ, Wareing DR, Maiden MC. Molecular characterization of *Campylobacter jejuni* clones: A basis for epidemiologic investigation. Emerging Infectious Diseases, 2002; 8:949–955.
14. Colles FM, Jones K, Harding RM, Maiden MC. Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment. Applied Environmental Microbiology, 2003; 69:7409–7413.
15. Kwan PS, Barrigas M, Bolton FJ, French NP, Gowland P, Kemp R, Leatherbarrow H, Upton MAJ, Fox AJ. The molecular epidemiology of *Campylobacter jejuni* populations in dairy cattle, wildlife and the environment in a farmland area. Applied Environmental Microbiology, 2008; 74:5130–5138.

16. Hald T. Quantifying Sources of Human Salmonellosis in the Netherlands and in Denmark: Two Approaches, Same Result? Danish Zoonosis Centre, Danish Veterinary Institute, 2002.
17. Wilson DJ, Gabriel E, Leatherbarrow AJ, Cheesbrough J, Gee S, Bolton E, Fox A, Fearnhead P, Hart CA, Diggle PJ. Tracing the source of campylobacteriosis. PLoS Genetics, 2008; 4:e1000203.
18. French N, Molecular Epidemiology and Veterinary Public Health Group Hopkirk Institute. Enhancing Surveillance of Potentially Foodborne Enteric Diseases in New Zealand: Human Campylobacteriosis in the Manawatu. New Zealand Food Safety Authority (NZFSA). Available at: http://www.nzfsa.govt.nz/science/research-projects/Campy_Attribution_Manawatu.pdf, Accessed on January 16, 2009.
19. Institute of Environmental Science and Research Limited (ESR). Public Health Surveillance—EpiSurv. Available at: http://www.surv.esrcri.nz/episurv/index/php, Accessed on January 16, 2009.
20. Toft N, Innocent GT, Gettinby G, Reid SWJ. Assessing the convergence of Markov chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard. Preventive Veterinary Medicine, 2007; 79:244–256.
21. Ben Aissa R, Al-Gallas N, Troudi H, Bethadj N, Belhadj A. Trends in *Salmonella enterica* serotypes isolated from human, food, animal, and environment in Tunisia, 1994–2004. Journal of Infection, 2007; 55:324–339.
22. Birk T, Rosenquist H, Brondsted L, Ingmer H, Bysted A, Christensen BB. A comparative study of two food model systems to test the survival of *Campylobacter jejuni* at −18 degrees C. Journal of Food Protection, 2006; 69:2635–2639.
23. Vora P, Senecal A, Schaffner DW. Survival of *Staphylococcus aureus* ATCC 13565 in intermediate moisture foods is highly variable. Risk Analysis, 2003; 23:229–236.
24. Evers EG, Van Der Fels-Klerx HJ, Nauta M, Schijven JF, Havelaar AH. *Campylobacter* source attribution by exposure assessment. International Journal of Risk Assessment and Management, 2008; 8:174–190.
25. French N, Mullner P. Risk-Based Tools to Inform Decision Making in Food Safety and Public Health—Applications in Europe and New Zealand FoodFrenz. Budapest, Hungary, 2008.
26. Mullner P, Spencer S, Collins-Emerson J, Midwinter A, Jones G, Noble A, French N. Molecular and Modeling Tools for *Campylobacter* Source Attribution. Proceedings of the Epidemiology and Animal Health Management Branch of the NZVA, 2008.
27. Wong TL, Hollis L, Cornelius A, Nicol C, Cook R, Hudson JA. Prevalence, numbers, and subtypes of *Campylobacter jejuni* and *Campylobacter coli* in uncooked retail meat samples. Journal of Food Protection, 2007; 70:566–573.
28. Oporto B, Esteban JI, Aduriz G, Juste RA, Hurtado A. Prevalence and strain diversity of thermophilic *campylobacters* in cattle sheep and swine farms. Journal of Applied Microbiology, 2007; 103:977–984.
29. Titus SM. A Novel Model Developed for Quantitative Microbial Risk Assessment in the Pork Food Chain. Institute of Veterinary, Animal and Biomedical Sciences, Massey University, 2007.
30. European Food Safety Authority. A Quantitative Microbiological Risk Assessment on *Salmonella* in Meat: Source Attribution for Human Salmonellosis from Meat. EFSA Panel on Biological Hazards, 2008.