Epidemiol. Infect. (2009), **137**, 897–905. © 2008 Cambridge University Press
doi:10.1017/S0950268808001209 Printed in the United Kingdom

Sources of sporadic *Yersinia enterocolitica* infection in children in Sweden, 2004: a case-control study

S. BOQVIST[1]*, H. PETTERSSON[2], Å. SVENSSON[3] AND Y. ANDERSSON[2]

[1] *Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Swedish Zoonosis Centre, National Veterinary Institute, Uppsala, Sweden*
[2] *Department of Epidemiology, Swedish Institute for Infectious Disease Control, Solna, Sweden*
[3] *Department of Mathematics, Stockholm University, Stockholm, Sweden*

(*Accepted 16 July 2008; first published online 15 September 2008*)

SUMMARY

Young children account for a large proportion of reported *Yersinia enterocolitica* infections in Sweden with a high incidence compared with other gastrointestinal infections, such as salmonellosis and campylobacteriosis. A case-control study was conducted to investigate selected risk factors for domestic sporadic yersiniosis in children aged 0–6 years in Sweden. In total, 117 cases and 339 controls were included in the study. To minimize exclusion of observations due to missing data a multiple non-parametric imputation technique was used. The following risk factors were identified in the multivariate analysis: eating food prepared from raw pork products (OR 3·0, 95% CI 1·8–5·1) or treated sausage (OR 1·9, 95% CI 1·1–3·3), use of a baby's dummy (OR 1·9, 95% CI 1·1–3·2) and contact with domestic animals (OR 2·0, 95% CI 1·2–3·4). We believe that the importance of *Y. enterocolitica* infection in children has been neglected and that results from this study can be used to develop preventive recommendations.

Key words: Case-control, children, foodborne illness, risk factor, *Yersinia enterocolitica*.

INTRODUCTION

Yersinia infection causes gastroenteritis and enterocolitis with symptoms of diarrhoea, vomiting, fever and abdominal pain [1]. In Sweden, infection with *Yersinia enterocolitica* is notifiable under the Swedish Communicable Disease Act and cases are reported by both physicians and laboratories to the Institute for Infectious Disease Control (SMI), Solna, Sweden, and to the county medical officers. The most commonly encountered bioserotype isolated from human cases in Sweden is *Y. enterocolitica* biotype 4 serotype O:3 which accounts for more than 95% of notified cases. There is a seasonal variation in number of reported cases with a peak from July to September. The incidence of reported domestically acquired yersiniosis caused by *Y. enterocolitica* during the year in which the present study was conducted (2004) was 9·0/100 000 inhabitants, as presented in the SMI database [2]. However, the incidence among children aged 0–6 years was 30/100 000 inhabitants. The SMI database also shows that 73% of reported cases in 2004 were domestically acquired and 28% were children aged <5 years. This can be compared with other notifiable gastrointestinal infections, such as salmonellosis and campylobacteriosis, for which children aged <5 years constituted 7% and 6%, respectively [2]. Compared

* Author for correspondence: S. Boqvist, DVM, Ph.D., Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, PO Box 7009, 750 07 Uppsala, Sweden.
(Email: Sofia.Boqvist@bvf.slu.se)

with these two infections, there is little information on the epidemiology of domestically acquired yersiniosis caused by *Y. enterocolitica*.

Out of previously published studies on factors associated with *Y. enterocolitica* infection, to our knowledge there is none that investigates the epidemiology of sporadic *Y. enterocolitica* infection in children. However, consumption of pork as a risk factor for *Y. enterocolitica* infection has been described in previous studies [3–5]. Other risk factors presented in these studies include more than two people living in the home [3] and eating sausage [5]. In a study conducted by Ostroff *et al*. [5], cases were also more likely to report a preference for eating meat prepared raw or rare and for drinking untreated water. Drinking water from a private well has also been associated with seropositivity to *Y. enterocolitica* [6]. These studies, together with other information [7, 8], show that pork is an important source of yersiniosis caused by *Y. enterocolitica*. Fattening pigs are important reservoirs of *Y. enterocolitica* and studies from Finland have shown *Y. enterocolitica* 4/O:3 to be present in 56% of tonsils of fattening pigs sampled at retail [9], 78% of tongues and in 2% of minced meat [10]. Furthermore, studies from Sweden [11] and Finland [12] show that *Y. enterocolitica* isolates from pork products and from patients with yersiniosis caused by the same pathogen are genetically indistinguishable.

In 2004, the Swedish National Food Administration (NFA) conducted a national survey to investigate, by PCR and culture, the presence of *Y. enterocolitica* in the most commonly consumed pork products in Sweden [13]. It was shown that DNA from *Y. enterocolitica* was present in 9% of the food products investigated.

The aim of our study was to investigate risk factors for *Y. enterocolitica* infection in children aged $\leqslant 6$ years in the products included in the study performed by the NFA, together with other potential risk factors.

METHODS

Study design

The study was a matched case-control study including all 21 counties in Sweden. Cases were defined as children aged $\leqslant 6$ years with a laboratory-confirmed *Y. enterocolitica* infection with onset of disease between 1 January and 31 December 2004. A reason for the chosen age cut-off was that children start compulsory school at age 7 years. Only cases with domestically acquired *Y. enterocolitica* infection, i.e. defined as the subject having been in Sweden within a 2-week period before onset of illness, were considered. When several cases were reported from the same family, only the first reported case was included. Cases reported >6 weeks after onset of illness were excluded due to risk of recall bias. Controls were selected based on three criteria: age (date of birth 1 month before or after the case), sex and place of residence, defined as the first three digits in the postal code of the case. The controls were selected from the computerized Swedish census database, which contains the social security numbers (including date of birth), name and addresses of all Swedish citizens. The observations listed closest (in age) above and below the case with the same sex and postal code were selected as controls.

Data collection

Data were collected from the parents of the children using a standardized postal questionnaire. Most questions had previously been used by the SMI in other questionnaires. For the cases, the questions related to the time period 2 weeks before onset of illness, while for the controls the time period was 2 weeks before receiving the questionnaire. A reminder and a new questionnaire were sent out 2 weeks after the first questionnaire in cases of no response.

The questionnaire included 28 questions relating to consumption of different food items, with special emphasis on pork products and other previously known, or suspected, food items as sources of infection. The pork products included were the same as in the NFA national study in which products consumed by all age groups or mainly by young children were selected [13]. The two main categories of pork products were raw and cold smoked/heat-treated products. In the present study, the food items were grouped together into 12 exposure groups based on the shared similarities (Table 1). Apart from this, other questions included in the analyses were contact with domestic animals, contact with different sources of outdoor water, food bought directly from a farm, use of a baby's dummy and playing in an outdoor sandbox. The questionnaire also contained questions about child daycare, housing, contact with other persons with gastrointestinal symptoms and source of drinking water, but these questions were excluded from further analyses due to

Table 1. *Description of variables investigated as potential risk factors in a case-control study of sporadic* Y. enterocolitica *infection in children aged 0–6 years in Sweden, 2004* (*percentages of imputed data and univariate odds ratio for the imputed and complete datasets are also shown*)

	% of imputed data		Univariate OR	
	Controls	Cases	Mean from five imputed datasets (range)	Complete dataset‡ (95% CI)
Food categories analysed (no. of grouped food items)				
Raw smoked pork, e.g. bacon (1)	3·8	21	0·62 (0·54–0·70)	0·82 (0·30–2·2)
Unprocessed raw pork products, e.g. pork chops (4)	6·8	25	2·5 (2·3–2·8)	3·0 (1·1–7·8)
Minced raw pork products, e.g. minced meat (2)	2·4	30	1·0 (0·88–1·2)	1·0 (0·32–3·3)
Ham ready-to-eat, i.e. smoked ham (3)	6·2	27	1·4 (1·2–1·5)	2·5 (0·86–7·6)
Sausage ready-to-eat, e.g. salami (3)	5·6	21	0·79 (0·68–0·90)	1·8 (0·63–5·3)
Sausage treated, e.g. Falun sausage* (2)	5·3	23	1·9 (1·8–2·1)	1·9 (0·69–5·5)
Other sausages, e.g. locally produced (2)	4·7	24	0·55 (0·53–0·57)	1·2 (0·48–3·0)
Paté ready-to-eat, e.g. brawn†, liver paté (4)	2·7	15	0·84 (0·77–0·90)	0·99 (0·41–2·4)
Kebab (possibly containing pork) (1)	1·2	10	0·34 (0·28–0·58)	1·7 (0·33–8·9)
Juice, fresh (1)	4·1	11	0·36 (0·34–0·38)	n.a.
Salad with mayonnaise, e.g. potato salad (4)	2·9	5·1	0·28 (0·27–0·29)	0·32 (0·04–2·5)
Unpasteurized milk	1·2	0	0·31 (0·29–0·31)	n.a.
Other variables analysed				
Use of a baby's dummy	0·03	0	1·8 (1·8–1·8)	1·3 (0·53–3·1)
Contact with animals	1·8	9·4	1·9 (1·8–2·0)	2·1 (0·74–5·9)
Playing in sandbox	4·7	5·1	1·1 (1·0–1·2)	1·2 (0·40–3·5)
Food bought from farmers' market/farm shop	5·6	7·7	0·47 (0·45–0·48)	n.a.
Contact with outdoor water source	3·5	15	0·45 (0·43–0·45)	n.a.

OR, Odds ratio; CI, confidence interval; n.a., not available.
* Sausage commonly consumed fried in Sweden.
† A dish made from processed pork traditionally consumed at Christmas.
‡ Questionnaires with no missing information ($n=232$).

the homogeneity of the responses. The cases were also questioned about symptoms, duration of illness, treatment and hospitalization.

The study was approved by the Ethics Committee of the Karolinska Institute, Stockholm, Sweden.

Statistics

Imputation

A substantial proportion of the questionnaires returned were only partially completed. Missing answers to particular questions consisted both of 'do not remember' and no response. Only 47% of the questionnaires were fully completed. When checking for patterns in the variables with missing data we found no influences of the characteristics of the child (age, gender, child daycare and housing). Furthermore, questionnaires from the cases had a larger number of partially missing data compared with controls, 81% and 42%, respectively (Fig. 1). The proportion of missing answers for the questions relating to specific food items varied between 12% and 35% for cases and 4% and 10% for controls.

After exploring the data it was decided to apply a multiple imputation technique to avoid excluding questionnaires with partially missing data. However, it was decided that if answers were missing on $\geqslant 10$ questions of the 28 questions about food intake, those questionnaires would be excluded in order not to bias the imputation. The imputation was done by taking into account age of the child and whether the child was a case or control. The reason for including age was that it is reasonable to believe that food intake differs with age. The five controls aged 6 years were grouped together with those aged 5 years, and cases aged 4, 5 and 6 years were also aggregated since there were few of them. This resulted in the use of 11 imputation groups, which contained 12–164 children.

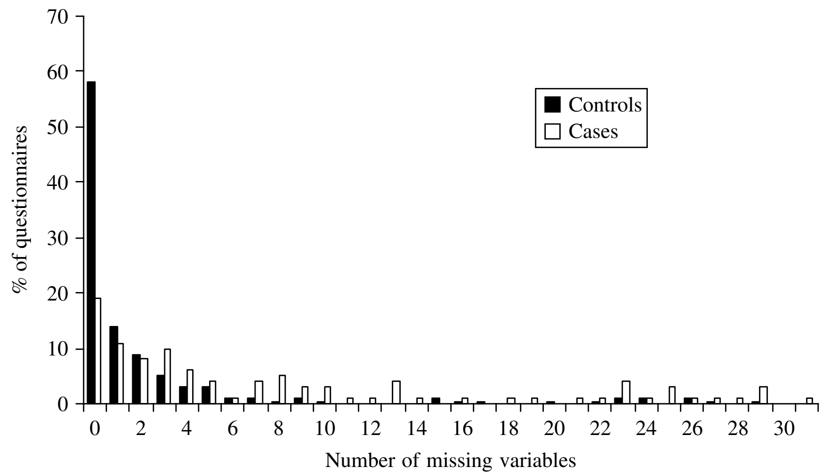

Fig. 1. Percentage of missing variables in a case-control study of sporadic *Y. enterocolitica* infection in children aged 0–6 years in Sweden, 2004.

To impute an answer in a specific questionnaire, the number of concordances between that questionnaire and all other questionnaires in the imputation group to which the specific questionnaire belonged were counted among the variables daycare and housing and the questions about food intake. For a specific questionnaire with a missing answer, the remaining questionnaires with the largest number of concordances N, $N-1$ and $N-2$ were selected as candidates to provide a possible imputed answer. From these candidates one was selected randomly, and its answer assigned to the specific questionnaire. This procedure was carried out five times with independent choices (with replacement) and thus generated five different datasets. The observed number of concordances, N, varied between 7 and 29 (mean number 19·3, median 20). The number of candidates, from which the values were imputed, varied between one (for eight persons) and 53 (mean number 12·6, median 11). This imputation procedure described above results in a non-parametric multiple imputation, i.e. no assumption is needed about the distribution of data. The percentage of imputations for each variable is presented in Table 1.

Statistical analysis

The controls were chosen to match the cases which make the design a matched case-control study. The matching was done according to the well-defined factors age, sex and postal code. We chose to analyse the data using logistic regression including age and sex as explanatory factors, rather than performing a conditional logistic regression analysis [14]. Age and sex were always included in the models in order to account for matching in the original design [14]. However, the estimates for their effects are not presented, since those would depend on the group matching procedures (and number of missing controls in different groups). In the logistic regression analysis case/control were used as outcome and age, sex and binary indicators of exposure groups (see section on data collection) were used as independent variables. Age was treated as a categorical variable in seven classes (0, 1, 2, 3, 4, 5, 6).

A child was regarded as belonging to a certain exposure group if it had been exposed to at least one of the food items in the group. If the answer was 'no' to all of the items in the group, the child was assumed to be unexposed.

Missing answers were imputed in the manner described above. We used multiple imputation whereby five imputed datasets were created and analysed separately. The five estimates of odds ratio (OR) for a single exposure group were pooled using methods described by Little & Rubin [15]. All variables (Table 1) were included in the model followed by backwards elimination, with a significance level of 5% to reduce the number of variables included. Three models were built, of which the first included all cases and controls. The other two models were built to investigate a potential difference between children aged 0–1 years (i.e. aged ⩽23 months) and 2–6 years, respectively. Logistic regression including one exposure group at a time based on the five imputed datasets and on complete-case data (i.e. the 232 observations with no missing data) were calculated.

χ^2 tests were used to compare symptoms for different age groups in the cases and a *P* value of $\leqslant 0{\cdot}05$ was considered significant.

RESULTS

Descriptive epidemiology

During the study period, questionnaires were sent to the parents of 205 cases that met the case definition and of those, 184 responded (90% response rate). Sixty-seven cases were excluded due to the following reasons: late notifications ($n=35$), travel abroad during the 2-week period before receiving the questionnaire ($n=2$), questionnaires being completed for the wrong child in the family ($n=7$), or information being missing for $\geqslant 10$ variables ($n=23$). Thus, 117 cases were included in total. The median time from onset of symptoms to response to the questionnaire was 34 days, ranging from 8 to 113 days.

The response rate for the controls was 67% (449/668), of which 339 were included. One hundred and ten controls were excluded due to: non-response of corresponding cases ($n=60$), travel abroad during the 2-week period before receiving the questionnaire ($n=30$), the questionnaire being completed for the wrong child in the family ($n=8$), or information being missing for $\geqslant 10$ variables ($n=12$). The median number of days from dispatch of control questionnaire to response was 8 days, ranging from 1 to 61 days.

Fifty per cent ($n=59$) of the cases were aged 1 year, followed by 20% ($n=23$) aged 2 years, 11% ($n=13$) aged <1 year, 9% ($n=11$) aged 3 years, 5% ($n=6$) aged 4 years, 3% ($n=3$) aged 5 years and 2% ($n=2$) aged 6 years. Of all cases, 60 were girls (51%) and 57 boys (49%). The frequency distributions for age and sex for the controls were similar. The most common reported symptom was diarrhoea, followed by stomach pain (Table 2). There were no statistical differences between symptoms reported for the cases aged 0–1 and 2–6 years, respectively. The median duration of symptoms was 16 days and 23 cases had been hospitalized for a medium duration of 3 days.

Odds ratios from univariate analyses of the imputed datasets and the complete dataset are presented in Table 1.

Table 2. *Clinical characteristics of 117 children aged 0–1 and 2–6 years included in a case-control study of sporadic* Y. enterocolitica *infection in children aged 0–6 years in Sweden, 2004*

	Cases with symptoms (%)	
Symptoms	Age 0–1 years	Age 2–6 years
Diarrhoea	70 (97)	44 (100)
Stomach pain	43 (88)	40 (98)
Fatigue	54 (86)	32 (80)
Fever	55 (83)	31 (74)
Body ache	9 (47)	12 (55)
Nausea	14 (48)	12 (41)
Vomiting	28 (43)	9 (24)
Headache	2 (13)	8 (40)
Joint pain	2 (14)	7 (33)
Dizziness	5 (25)	2 (9)
Other symptoms	10 (29)	6 (21)

Risk factors

For children aged between 0 and 6 years, foods prepared from unprocessed raw pork products (e.g. pork chop: OR 3·0, 95% CI 1·8–5·1) and treated sausages (e.g. salami: OR 1·9, 95% CI 1·1–3·3) were associated with domestically acquired *Y. enterocolitica* infection (Table 3). Other factors associated with infection were use of a baby's dummy (OR 1·9, 95% CI 1·1–3·2) and contact with domestic animals (OR 2·0, 95% CI 1·2–3·4). Eighty-two cases reported contact with domestic animals, of which 81 were dogs or cats.

In the next analyses two age groups were established to compare potential age differences. For children aged 0–1 years the following risk factors were associated with *Y. enterocolitica* infection: consumption of food prepared from unprocessed raw pork products (OR 2·7, 95% CI 1·3–5·4) and treated sausages (OR 2·6, 95% CI 1·2–5·5), as well as use of a baby's dummy (OR 2·0, 95% CI 1·0–4·0; Table 4). For children aged between 2 and 6 years the following factors were associated with *Y. enterocolitica* infection: consumption of food prepared from raw pork products (OR 4·4, 95% CI 1·7–11·3) and contact with animals (OR 3·1, 95% CI 1·2–8·3; Table 4).

DISCUSSION

The present study identified risk factors for sporadic yersiniosis caused by *Y. enterocolitica* in Sweden and to the authors' knowledge, this is the first case-control study of *Y. enterocolitica* infection in children. The majority of reported *Y. enterocolitica* infections in

https://doi.org/10.1017/S0950268808001209 Published online by Cambridge University Press

Table 3. *Results from multivariate logistic regression in a case-control study of sporadic domestically acquired* Y. enterocolitica *infection in children aged 0–6 years in Sweden, 2004*

Risk factor (% imputed missing factors)	Cases (%) ($n=117$)	Controls (%) ($n=339$)	Multivariate OR (95% CI)
Food categories			
Raw smoked pork (bacon) (8)	23 (20)	98 (29)	0·5 (0·2–0·9)
Unprocessed raw pork products (11)	66 (56)	140 (41)	**3·0 (1·8–5·1)**
Juice (6)	11 (9)	80 (24)	0·5 (0·2–0·9)
Other sausages (10)	38 (32)	165 (49)	0·6 (0·4–0·97)
Salad with mayonnaise (4)	5 (4)	47 (14)	0·3 (0·1–0·7)
Sausage treated (10)	77 (66)	190 (56)	**1·9 (1·1–3·3)**
Other			
Use of a baby's dummy (0·2)	76 (65)	174 (51)	**1·9 (1·1–3·2)**
Contact with animals (4)	82 (70)	223 (66)	**2·0 (1·2–3·4)**

OR, Odds ratio; CI, confidence interval.
Significant values indicated in bold.

Table 4. *Results from multivariate logistic regression in a case-control study of sporadically acquired domestic* Y. enterocolitica *infection in children aged 0–1 and 2–6 years in Sweden, 2004*

Risk factor (% imputed missing factors)	Cases (%)	Controls (%)	Multivariate OR (95% CI)
Age group 0–1 years	($n=72$)	($n=180$)	
Food categories			
Raw smoked pork (bacon) (11)	49 (68)	127 (71)	0·3 (0·1–0·97)
Unprocessed raw pork products (11)	31 (43)	58 (32)	**2·7 (1·3–5·4)**
Other sausages (9)	19 (26)	81 (45)	0·4 (0·2–0·9)
Salad with mayonnaise (4)	2 (3)	29 (16)	0·1 (0·02–0·6)
Sausage treated (9)	44 (61)	90 (50)	**2·6 (1·2–5·5)**
Other			
Use of baby's dummy (0)	56 (78)	115 (64)	**2·0 (1·0–4·0)**
Age group 2–6 years	($n=45$)	($n=159$)	
Food categories			
Unprocessed raw pork products (12)	35 (78)	82 (52)	**4·4 (1·7–11·3)**
Other			
Playing in sandbox (7)	34 (76)	140 (88)	0·4 (0·1–0·97)
Contact with animals (4)	39 (87)	108 (68)	**3·1 (1·2–8·3)**

OR, Odds ratio; CI, confidence interval.
Significant values indicated in bold.

Europe are sporadic, although there have been recorded cases of foodborne outbreaks caused by a wide variety of food items, such as milk [16, 17], chocolate milk [18], soybean curd [19], salami [20] and brawn (a dish made from processed pork traditionally consumed at Christmas [21, 22]).

The risk factors investigated in the study mainly included pork or products thereof, but also other factors that have previously been suspected to be sources of actual *Y. enterocolitica* infections or potential sources of infection. The reason for the special emphasis on pork products was because the national survey was conducted by the NFA to investigate the presence of *Y. enterocolitica* in pork products during the same year in which the present study was conducted [13]. Pigs are considered the main reservoir for *Y. enterocolitica* [7, 8] and it can therefore be suspected that consumption of food prepared from pork would constitute an important source of sporadic *Y. enterocolitica* infection. This was confirmed in the present study, in which pork products and treated sausages, probably containing pork, were identified as risk factors. These associations have also been reported previously in the literature [3–5] and studies from the Nordic countries have shown genetically identical *Y. enterocolitica* isolates from pork products and from patients with yersiniosis [11, 12]. Cross-contamination of treated food products has previously been a suspected source of a foodborne outbreak caused by *Y. enterocolitica* in Sweden [21] and might be one reason for the association found in our study with pre-boiled or heat-treated sausages. Furthermore, mistakes in the production chain may also lead to products containing viable *Y. enterocolitica*. In the NFA study, DNA from *Y. enterocolitica* was found in products of untreated (10%) and treated pork (7% [13]). This shows that viable *Y. enterocolitica* were probably present in the food products, thereby supporting results found in our study.

Contact with domestic animals was also identified as a risk factor. A previous study has shown that raw pork can be a source of infection in dogs and cats [23] and dogs may excrete *Y. enterocolitica* in their faeces for weeks [24]. This supports the finding that companion animals may be sources of human infection. In the present study, almost 70% of the cases reported contact with dogs or cats, which reflects the popularity of cats and dogs as companion animals. Children having contact with animals should preferably be under the supervision of an adult.

Separate statistical analyses were performed for children aged 0–1 and 2–6 years, respectively. The reason for this was that it could not be assumed that these age groups would have the same pattern of exposure and also because almost 50% of the cases were children aged 1 year. It can be speculated that parents of younger children usually bring their ill child to medical care more often than parents of children in the older age groups. However, there was no difference in clinical signs between the two age groups. While consumption of food prepared from unprocessed raw pork was a risk factor for both age groups consumption of treated sausage was a risk factor only for children aged 0–1 year. One explanation might be that the infectious dose might be lower for the younger children and that children aged 1 year often start eating food consumed by older children and adults, which might lead to exposure to new pathogens.

One reason why contact with animals was only a risk factor for the older age group and use of a baby's dummy for the younger group might be due to different exposure patterns. It may be that older children are often more interested in having close contact with companion animals, for example feeding animals from the hand and petting, whereas younger children usually use a baby's dummy more often. It seems likely that a baby's dummy can be a route of transmission if it becomes contaminated by being dropped onto the ground.

In the multivariate analyses some factors negatively associated with *Y. enterocolitica* infection were identified (Tables 3 and 4). It was decided to keep them in the model as their inclusion improved the general fit of the model.

In our study the controls were not questioned about the same time period as the cases. It seems likely that the parents of cases would be more likely to remember different food items consumed before onset of symptoms, but it might also be that it would be more difficult for the parents of cases to remember food consumed up to 6 weeks prior to answering the questionnaire, compared with parents of controls. An effect of this was that controls were more likely to remember exposure to the potential risk factors than cases. For some cases and controls the information provided was mainly about food consumed in the home, as parents reported that they did not know what food their children had eaten at daycare. It can also be discussed whether the questionnaires completed a number of weeks after onset of illness may have given information about eating habits rather than food consumed during the 2-week period before illness. There was also a difference in time period between onset of illness for the cases/dispatch of questionnaire to controls and date of completion of the questionnaire. This difference is not possible to circumvent completely as there is a delay in laboratory analysis and in the reporting system and as cases cannot be contacted before information about the infection has been submitted by their physician. This is probably a concern in most case-control studies. Thus, there was probably a risk of recall bias in the study.

Seasonal variations in number of reported cases have been recorded for different gastrointestinal infections, including yersiniosis [2]. Explanations for this might be increased outdoor activities and changes in food consumption during summer leading to different exposure patterns. However, seasonality was not investigated for in the present study.

Although the response rate was high, a high percentage of questionnaires had partially missing data. This is commonly seen when questionnaires are sent by post to respondents. The response rate to individual questions is often better if a telephone interview is conducted. The majority of published case-control studies exclude observations with missing data, which can lead to a great loss of information for those variables where data are present. If that strategy had been used in the present study, a great number of observations would have been lost, thereby reducing important information and possibly biasing the results of the statistical analysis. Missing data were partly dealt with by grouping together food items that were similar, whereby the answer 'yes' to one food item resulted in 'yes' for the group of food items. Investigation of missing data showed that there were more missing data in the food categories commonly consumed by Swedish families with children (e.g. minced meat and ham) compared with food categories consumed less often (e.g. salads with mayonnaise). It is possibly easier to specifically remember food consumed on special occasions or more rarely. Another way to avoid loss of information was use of imputation to estimate data for the missing variables. However, there is always a risk of imputation adding bias. In our study, the imputed datasets and the complete dataset (i.e. the cases or controls with no missing answers) were analysed. The comparison showed that the estimates of the odds ratios from the imputed datasets were closer to 1 and had narrower confidence intervals than the estimates from the complete set. There were small differences in the results obtained from the five different imputed datasets, which implies that the actual imputation did not cause much additional variation. Another measure used to reduce bias was to exclude observations with $\geqslant 10$ missing variables so as not to make too many imputations per observation.

To reduce the incidence of *Y. enterocolitica* infection in children in Sweden special attention should be paid to children aged 1 year. It is important to reduce the number of viable *Y. enterocolitica* in pork products, or products containing pork, by ensuring correct handling at production, retail and consumer level. Special attention should also be paid to potential sources of infection in the surrounding environment, for example to avoid contamination of babies' dummies. Further investigations are needed to determine why children aged 1 year are overrepresented in reported domestically acquired sporadic cases of *Y. enterocolitica* infection.

ACKNOWLEDGEMENTS

The authors are most grateful to the parents of all cases and controls who contributed information. Without their help it would not have been possible to conduct this study.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Cover TL, Aber RC.** *Yersinia enterocolitica. New England Journal of Medicine* 1989; **321**: 16–24.
2. **SMI database** (http://www.smittskyddsinstitutet.se/statistik/yersiniainfektion/). Accessed 15 April 2007.
3. **Satterthwaite P, *et al.*** A case-control study of *Yersinia enterocolitica* infections in Auckland. *Australian and New Zealand Journal of Public Health* 1999; **23**: 482–485.
4. **Tauxe RV, *et al.*** *Yersinia enterocolitica* infections and pork: the missing link. *Lancet* 1987; **329**: 1129–1132.
5. **Ostroff SM, *et al.*** Sources of sporadic *Yersinia enterocolitica* infections in Norway: a prospective case-control study. *Epidemiology and Infection* 1994; **112**: 133–141.
6. **Saebo A, *et al.*** Prevalence of antibodies to *Yersinia enterocolitica* O:3 among Norwegian military recruits: association with risk factors and clinical manifestations. *European Journal of Epidemiology* 1994; **10**: 749–755.
7. **Kapperud, G.** *Yersinia enterocolitica* in food hygiene. *International Journal of Food Microbiology* 1991; **12**: 53–65.
8. **Bottone EJ.** *Yersinia enterocolitica*: overview and epidemiologic correlates. *Microbes and Infection* 1999; **1**: 323–333.
9. **Korte T, *et al.*** Low prevalence of yadA-positive *Yersinia enterocolitica* in sows. *Foodborne Pathogens and Disease* 2004; **1**: 45–52.
10. **Fredriksson-Ahomaa M, Hielm S, Korkeala H.** High prevalence of yadA-positive *Yersinia enterocolitica* in pig tongues and minced meat at the retail level in Finland. *Journal of Food Protection* 1999; **62**: 123–127.
11. **Thisted Lambertz S, Danielsson-Tham M-L.** Identification and characterization of pathogenic *Yersinia enterocolitica* isolates by PCR and pulsed-field gel electrophoresis. *Applied and Environmental Microbiology* 2005; **71**: 3674–3681.
12. **Fredriksson-Ahomaa M, *et al.*** Correspondence of genotypes of sporadic *Yersinia enterocolitica* bioserotype 4/O:3 strains from human and porcine sources. *Epidemiology and Infection* 2001; **127**: 37–47.
13. **Thisted Lambertz S.** Pathogenic *Yersinia enterocolitica* in non-treated and treated pork products [in Swedish]. Riksprojekt 2004. Livsmedelsverkets rapport no. 18/2005, 2005, pp. 42.

14. **Clayton D, Hills M.** *Statistical Models in Epidemiology*. New York: Oxford University Press, 1993, pp. 182.
15. **Little RJA, Rubin DB.** *Statistical Analysis with Missing Data*, 2nd edn. New York: John Wiley, 2002, pp. 86–87.
16. **Ackers M, *et al.*** An outbreak of *Yersinia enterocolitica* O:8 infections associated with pasteurized milk. *Journal of Infectious Diseases* 2000; **181**: 1834–1837.
17. **Tacket CO, *et al.*** A multistate outbreak of infections caused by *Yersinia enterocolitica* transmitted by pasteurized milk. *Journal of the American Medical Association* 1984; **251**: 483–486.
18. **Black RE, *et al.*** Epidemic *Yersinia enterocolitica* infection due to contaminated chocolate milk. *New England Journal of Medicine* 1978; **298**: 76–79.
19. **Tacket C, *et al.*** An outbreak of *Yersinia enterocolitica* infections caused by contaminated tofu (soybean curd). *American Journal of Epidemiology* 1985; **121**: 705–711.
20. **Harb J, *et al.*** Case-control study assessing the association between yersiniosis and exposure to salami. *Canada Communicable Disease Report* 2000; **26**: 161–164.
21. **SMI database** (http://www.smittskyddsinstitutet.se/sjukdomar/yersiniainfektion/) Accessed 15 April 2007.
22. **Grahek-Ogden D, Nygård K, Lassen J.** Outbreak of *Yersinia enterocolitica* O:9 infections in Vestfold and Ostfold [in Norwegian]. *MSIS-rapport, Meldinssystem for smittsomme sykdommer*. Folkehelseinstituttet, Norway, 2006; **34**, 7.
23. **Fredriksson-Ahomaa M, Korte T, Korkeala H.** Transmission of *Yersinia enterocolitica* 4/O:3 to pets via contaminated pork. *Letters in Applied Microbiology* 2001; **32**: 375–378.
24. **Fredriksson-Ahomaa M, Stolle A, Korkeala H.** Molecular epidemiology of *Yersinia enterocolitica* infections. *FEMS Immunology & Medical Microbiology* 2006; **47**, 315–329.

https://doi.org/10.1017/S0950268808001209 Published online by Cambridge University Press