Attributing sporadic and outbreak-associated infections to sources: blending epidemiological data

D. COLE[1]*, P. M. GRIFFIN[1], K. E. FULLERTON[1], T. AYERS[2], K. SMITH[3], L. A. INGRAM[4], B. KISSLER[5] AND R. M. HOEKSTRA[2]

[1] *Enteric Diseases Epidemiology Branch, Division of Foodborne, Waterborne, and Environmental Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA*
[2] *Biostatistics and Information Management Office, Division of Foodborne, Waterborne, and Environmental Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA*
[3] *Foodborne, Vectorborne, and Zoonotic Diseases Unit, Acute Disease Investigation and Control Section, Minnesota Department of Health, St Paul, MN, USA*
[4] *Communicable and Environmental Diseases and Emergency Preparedness, Tennessee Department of Health, Nashville, TN, USA*
[5] *Applied Epidemiology Division, Office of Public Health Science, Food Safety and Inspection Service, US Department of Agriculture, Atlanta, GA, USA*

SUMMARY

Common sources of shiga toxin-producing *Escherichia coli* (STEC) O157 infection have been identified by investigating outbreaks and by case-control studies of sporadic infections. We conducted an analysis to attribute STEC O157 infections ascertained in 1996 and 1999 by the Foodborne Diseases Active Surveillance Network (FoodNet) to sources. Multivariable models from two case-control studies conducted in FoodNet and outbreak investigations that occurred during the study years were used to calculate the annual number of infections attributable to six sources. Using the results of the outbreak investigations alone, 27% and 15% of infections were attributed to a source in 1996 and 1999, respectively. Combining information from both data sources, 65% of infections in 1996 and 34% of infections in 1999 were attributed. The results suggest that methods to incorporate data from multiple surveillance systems and over several years are needed to improve estimation of the number of illnesses attributable to exposure sources.

Key words: Epidemiology, *Escherichia coli*, foodborne infections.

INTRODUCTION

Shiga toxin-producing Escherichia coli (STEC) O157 is estimated to cause 96 534 illnesses each year in the USA; domestically acquired, foodborne infections comprise 63 000 of these, with 2000 hospitalizations and 20 deaths annually [1]. Estimating the number of illnesses caused by pathogens commonly transmitted by food is an important step in determining public health priorities. Assessing the proportions of these illnesses due to specific exposures contributes to the development of targeted disease prevention strategies and helps to measure progress towards food safety goals.

Estimating the proportion of illnesses that can be attributed to contaminated foods is complicated by

* Author for correspondence: Dr D. Cole, DVM, PhD, Lead, Analytics Team, Enteric Diseases Epidemiology Branch, Division of Foodborne, Waterborne and Environmental Diseases, 1600 Clifton Rd, NE, MS C-09, Atlanta, GA 30333, USA.
(Email: dcole@cdc.gov)

the fact that foodborne pathogens can also be transmitted through a variety of other exposure pathways, such as direct contact with animals or with other infected persons, and by exposure to contaminated water. Outbreak investigations represent a unique opportunity to directly attribute illnesses to specific exposure sources. Many sources of STEC O157 infection have been identified in outbreak investigations, including ground beef, lettuce, juice, sprouts, spinach, and contact with animal manure or contaminated water [2–4]. Information from outbreak surveillance was used to estimate that 68% of domestically acquired STEC O157 infections are foodborne [1]. However, <30% of STEC O157 infections are associated with recognized outbreaks [5, 6]. To determine sources of STEC O157 infection in cases that are not epidemiologically linked to outbreaks (i.e. sporadic infections), FoodNet has conducted case-control studies; these studies have been used to calculate the population attributable fractions (PAFs) associated with specific exposures [7, 8].

The use of attributable fractions as estimated by case-control studies to determine the proportions of infections that could be prevented if specific exposures were removed is well described [9]. The FoodNet surveillance population is particularly well suited for calculating PAFs because active laboratory surveillance to identify incident infections supports the assumption that the source population accurately represents the target FoodNet catchment population [10, 11]. Therefore, the FoodNet surveillance system provides a unique opportunity to attribute both outbreak-associated and sporadic infections to specific exposure sources and to develop a more complete picture of attribution within a population. This study uses the results of two case-control studies of sporadic STEC O157 infections conducted in FoodNet in 1996 and 1999 and the outbreak investigations that occurred during the two study periods to attribute infections to specific exposure sources.

METHODS

Data sources

FoodNet is a collaborative programme among CDC, 10 state health departments, the U.S. Department of Agriculture's Food Safety and Inspection Service, and the Food and Drug Administration. It has conducted active, population-based surveillance for laboratory-confirmed cases of infection caused by STEC O157 since 1996, and is part of CDC's Emerging Infections Program. FoodNet personnel regularly contact clinical laboratories to ascertain laboratory-confirmed cases of infection occurring within the surveillance sites [12]. In 1996 and 1999, FoodNet sites conducted year-long case-control studies of sporadic STEC O157 infections [7, 8]. Data were obtained from these studies and from outbreak investigations conducted in FoodNet sites during these study years.

Outbreak investigations

Outbreaks are defined as incidents in which two or more persons experience a similar illness resulting from exposure to a common source. Personnel detect outbreaks in FoodNet sites in a variety of ways, most commonly by reports from private citizens and medical professionals, and by reports of notifiable diseases from clinical laboratories [13]. Investigation of outbreaks often includes active case-finding and the collection of stool specimens for identification of the aetiological agent. Outbreak cases can be defined by a common exposure setting even when a specific contaminated source cannot be implicated by outbreak investigators. Summary information regarding the exposure setting and source of infections for reported outbreak cases was obtained directly from the FoodNet sites for this study. Only laboratory-confirmed cases were included in the analysis.

Case-control studies

Case-control studies for both years used the same patient and control enrolment methodologies and criteria for eligibility, exclusion, and matching [8]. All laboratory-confirmed STEC O157 infections were ascertained in the FoodNet catchment populations through active surveillance, and cases were interviewed within 21 days of their stool sample collection date about exposures in the 5 days [7] or 7 days [8] before illness onset. Sequential digit dialling and birth registries were used to ascertain age-matched controls. Age- and telephone area code-matched controls were interviewed within 7 days of case-patient interviews. Potential controls were excluded if they reported having diarrhoea within 28 days prior to the case-patient's illness onset. In both case-control studies, infections associated with confirmed outbreaks (i.e. infections that were epidemiologically linked to exposure sources) were excluded and

described separately [7, 8]. During both studies, interviews with cases and controls included questions about travel, child daycare, exposure to farms and farm animals, use of appropriate food handling practices, exposure to untreated water sources, settings of food consumption, and consumption of meats (including different cuts of beef), fruits, and vegetables.

Data analysis

For each study year, sporadic STEC O157 infections were attributed only to those exposures that were significantly and positively associated with infection in the final multivariable model in the published paper [7, 8]. When the exposure variable assessed did not specify a contaminated source, the most likely source associated with that exposure was assumed (e.g. contaminated food for the exposure 'ate at a table-service restaurant', and animal contact for the exposure 'living on, working, or visiting a cattle farm'). Sporadic cases were not attributed to exposures that were associated with susceptibility to infection (e.g. use of immunosuppressive medication). For each outbreak, all associated infections were attributed to the specific source implicated by outbreak investigators. Three cases were identified to be outbreak-related, but a common exposure source was not reported so these were excluded from analysis.

Both case-control studies were conducted in dynamic populations, and controls were enrolled simultaneously with cases so they were at risk of infection during the same time as cases. Consequently, the calculated odds ratios are considered adequate estimators of the rate ratios or relative risks in the target population [14], and these were used to calculate the attributable fraction for each exposure source using the following formula [9–11]:

$$\text{AF}_{\text{pop}} = (\text{pd}_{\text{exp}}) * [(\text{OR}_{\text{exp}} - 1)/\text{OR}_{\text{exp}}],$$

where the PAF (AF_{pop}) for a specific exposure is a function of the proportion of cases exposed (pd_{exp}) to a specific source and the adjusted odds ratio for that exposure (OR_{exp}). The number of sporadic infections in the FoodNet study population attributable to each exposure during the study year was calculated by the following formula:

$$\text{AN}_{\text{exp}} = (\text{AF}_{\text{pop}}) * (\text{total no. sporadic cases}),$$

where the number of sporadic cases attributable to a specific exposure (AN_{exp}) is determined by the AF_{pop} and the total number of sporadic cases included in the case-control study [9–11]. After calculating the

number of sporadic infections attributable to each exposure source, the number of unattributed infections was determined by subtracting the sum of all attributed cases from the total number of cases included in the case-control study. Attribution of sporadic infections not included in the case-control studies was done by applying the same AF_{pop} to the total number of excluded sporadic cases.

Sources of uncertainty associated with the attribution method

To generalize the results of estimated attributable fractions to the target population (the FoodNet surveillance population), two important assumptions were needed; the validity of these are uncertain. First, the source population of laboratory-confirmed STEC O157 infections was assumed to be representative of the target population (FoodNet catchment) so that the attributed sources of infection in the source population of cases reflected those in the target population [11]. For this to be valid, it was assumed that there were no underlying differences in the distribution of causal transmission pathways in the three subpopulations of STEC O157 cases: (1) sporadic cases enrolled in each case-control study, (2) sporadic cases not enrolled in the case-control studies, and (3) outbreak-associated cases. It was also assumed that all laboratory-confirmed sporadic and outbreak-associated infections occurring in the surveillance population during the two study years were ascertained (i.e. sampling fraction = 1). In addition, all outbreak-associated illnesses were assumed to be due to the sources implicated in the outbreak investigations. Likewise, it was assumed that all ascertained infections were correctly classified as sporadic or outbreak-associated such that the distribution of exposures in the source population of sporadic cases were not influenced by unrecognized outbreaks in cases included in the case-control studies.

Second, the exposure risks identified in the study population and used to calculate the attributable fractions were assumed to reflect the effect of exposure in the target population [11]. This means that survey questions and categories of exposure used in both the case-control studies and outbreak investigations reflected relevant exposure definitions of infection risk [15, 16]. Likewise, it was assumed that confounding was adjusted for in the multivariable models [10, 11, 17], and that there was no interaction between exposure factors [9, 10]. It is important to note that

the sum of AN_{exp} in a given study year could exceed the total number of sporadic cases ascertained because a case could be counted more than once if exposed to more than one significant exposure [10, 18].

RESULTS

FoodNet sites ascertained 396 and 531 STEC O157 cases in 1996 and 1999, respectively. One hundred and five cases were associated with seven outbreaks in 1996. Seventy-nine cases were associated with four outbreaks in 1999. Table 1 provides the attributable numbers for the 740 sporadic cases and 184 outbreak-associated cases ascertained in the two study years. Of the 479 cases enrolled in the two case-control studies (68% and 63% of reported sporadic cases in 1996 and 1999, respectively), 53% (1996) and 23% (1999) were attributed to a source. After attribution fractions derived from cases enrolled in the case-control studies were applied to sporadic cases not enrolled, and this information was added to the case-control study results and outbreak investigations, attribution to exposure sources increased to 65% of all cases ascertained in 1996 and 34% of all cases ascertained in 1999.

Thirty-four per cent of sporadic cases were attributed to foodborne exposures in 1996 and 10% in 1999. Foodborne exposures accounted for 26% of outbreak cases in 1996 and 18% in 1999. Foodborne exposures in 1996 included exposure at restaurants or other food service settings to an undetermined source, presumably food (20% of sporadic and 7% of outbreak cases), hamburger (14% of sporadic and 2% of outbreak cases), and produce (17% of outbreak cases, specifically apple cider and lettuce). Food exposures in 1999 were hamburger (10% of sporadic and 14% of outbreak cases) and produce (4% of outbreak cases, all linked to romaine lettuce).

In 1996, attributed non-foodborne exposure sources were contact with animals (13% of sporadic cases), exposure to untreated water (4% of outbreak cases), and person-to-person transmission (6% of sporadic cases and 70% of outbreak cases). In 1999, attributed non-foodborne exposures were contact with animals (8% of sporadic cases), exposure to untreated water (5% of sporadic and 73% of outbreak cases), and person-to-person transmission (9% of outbreak cases).

DISCUSSION

This study shows that applying the information obtained from case-control studies to other sporadic infections, and adding these to data obtained from outbreak investigations, can increase the number of infections that are attributed to a source. Source attribution using data from outbreak surveillance provides the most definitive links to sources because an outbreak is, by definition, characterized by cases that share a causal pathway. However, most infections are sporadic; case-control studies are used to identify differences between the exposure profiles of infected persons and a control population to determine the attributable fractions associated with exposures found to be statistically significant in a multivariable model. Sources of sporadic infection include a wide variety of transmission pathways, each of which may have particular factors (e.g. food storage or preparation practices) contributing to individual illnesses that cannot be quantitatively captured by the study [19]. Consequently, the number of infections attributed to specific sources by case-control studies is typically small. The attribution of a higher proportion of ascertained cases in the population when data from both sporadic and outbreak-associated infections are used shows the value of combining two types of information to estimate sources of illness in a population. However, interpretation of the differences in the proportions of STEC O157 infections attributable to specific exposures when using blended estimates compared to those derived from only one data source, as well as the differences in attribution between the two study years, is complicated by two important sources of uncertainty: (a) the validity of the assumptions associated with the method used to attribute infections to exposure sources, and (b) the inherent variability in source attribution over time.

The assumption that there are no underlying differences in the distribution of causal transmission pathways of infection in the three subpopulations of laboratory-confirmed STEC O157 cases is difficult to evaluate. This is because exposure sources are determined indirectly for sporadic infections but directly for outbreak-associated infections. Use of active surveillance to increase the sampling fraction of source cases (laboratory-confirmed STEC O157 infections) helps to reduce this uncertainty by increasing the probability that the study population reflects the source population [11]. However, it is known that laboratory-confirmed infections represent a subset of the total number of illnesses in the population [1], and the degree to which confirmed infections represent the distribution of exposures and illness in the target population is unknown. In addition, sporadic cases who reported

Table 1. *Number of STEC O157 cases (and proportion of row total) ascertained by FoodNet Surveillance in 1996 and 1999 attributed to exposure sources identified in case-control studies and outbreaks investigated during the study period*

| | | Number of cases (row %) | | | | | | | |
| | | Foodborne exposures | | | Non-foodborne exposures | | | Other/ unknown exposures | |
	Study year	Hamburger	Produce	Food service	Animal contact[a]	Untreated water	Person-to-person		Total
Enrolled in case-control study	1996	28 (14%)[b]	0	38 (20%)[c]	26 (13%)	0	11 (6%)[d]	93 (47%)	196
	1999	29 (10%)[e]	0	0	22 (8%)	14 (5%)[f]	0	218 (77%)	283
Not enrolled in case-control study[g]	1996	13 (14%)	0	18 (20%)	12 (13%)	0	5 (6%)	44 (47%)	92
	1999	17 (10%)	0	0	13 (8%)	8 (5%)	0	131 (77%)	169
Part of recognized outbreaks	1996	2 (2%)	18 (17%)[h]	7 (7%)	0	4 (4%)[i]	74 (70%)[j]	0 (0%)	105
	1999	11 (14%)	3 (4%)[k]	0	0	58 (73%)[l]	7 (9%)[j]	0 (0%)	79
Exposure source	1996	43 (11%)	18 (4%)	63 (16%)	38 (10%)	4 (1%)	90 (23%)	140 (36%)	393
Total	1999	57 (11%)	3 (0·6%)	0	35 (7%)	80 (15%)	7 (1%)	349 (66%)	531

[a] Estimated by exposures 'visited farm with cows' (1996), 'lived on farm or visited farm' (1996), and 'living on, working, or visiting a cattle farm' (1999).
[b] Estimated by exposures 'eating pink hamburger at home' and 'eating pink hamburger away from home'.
[c] Estimated by exposure 'table-service restaurant'.
[d] Estimated by exposure 'child less than 2 years of age in household'.
[e] Estimated by exposure 'pink hamburger'.
[f] Estimated by exposure 'drinking untreated surface water'.
[g] It was assumed that exposure distribution in unenrolled cases was the same as enrolled cases.
[h] Due to apple cider (10 cases) and lettuce (8 cases).
[i] Due to swimming in a lake.
[j] Due to transmission in a childcare centre.
[k] Due to romaine lettuce.
[l] Due to consumption of well water contaminated by cattle faeces.

an ill household member within 28 days before submission of their stool sample were excluded from the case-control study. It is likely that a higher proportion of these excluded cases were actually related to person-to-person transmission compared to those enrolled in the study, but we assumed that both populations of sporadic cases experienced the same distribution of risky exposures. Furthermore, during at least one outbreak investigation included in this analysis, investigators conducted additional laboratory testing (Minnesota, 1996), resulting in additional laboratory-confirmed infections being identified in persons who had not sought healthcare. As a result, cases with outbreak-related illness were over-sampled relative to those associated with sporadic infection. These observed underlying differences in the ascertainment of infections from sporadic and outbreak-associated subpopulations illustrate key differences in the surveillance populations that probably influence the observed distributions of exposures and resulting attribution estimates, and may limit the generalizability of results to the target population.

Another assumption used to attribute infections to sources is that cases were correctly classified as sporadic or outbreak-associated, such that PAFs obtained from the case-control studies represent valid estimates of attribution. For example, in the 1996 study, the independent exposures of eating at a table service restaurant and eating 'pink' hamburger away from home were associated with PAFs of 20% and 7%, respectively. However, neither of these exposures was significantly associated with infection in the 1999 study. Under the assumption that all infections were correctly classified as sporadic or outbreak-associated during both study years, these differences in PAF reflect changes in the attributable fractions associated with these exposures between the two study years. However, if some of the sporadic cases reporting these two exposures in 1996 were associated with an undetected outbreak caused by contaminated ground beef (or exposure to ingredients cross-contaminated by ground beef during food preparation) distributed to restaurants, the exposure frequencies of these two variables in study cases may have been increased by this outbreak. In this scenario, detection of the outbreak would have resulted in the exclusion of these cases from the study, and these two exposures may not have been retained in the 1996 multivariable model, resulting in more similar final models and cumulative attribution estimates across the two case-control studies.

The exclusion of three cases from this analysis highlights the difficulty associated with assigning all cases as either sporadic or outbreak-associated. Although the reporting state designated these three cases as outbreak-associated, no common exposure was reported. We were unable to determine whether a common source was identified but not reported, or whether the cases had a common venue, e.g. eating at the same restaurant, but the specific exposure was not known. In the past decade, the Molecular Subtyping Network for Foodborne Disease Surveillance (PulseNet) has improved our ability to identify small numbers of possibly related cases, e.g. with the same rare subtype pattern, clustered in time and space. It is likely that many of these clusters represent common source outbreaks, but successful identification of a single exposure source in small numbers of clustered infections is difficult and relatively uncommon [20]. Therefore, these clustered cases are classified as sporadic and potentially influence PAFs estimated by case-control studies.

The validity of the assumption that exposure risks identified in the study population reflect the effect of exposures in the target population is highly uncertain in case-control studies. Ascertained exposures in case-control studies are assumed to be descriptive of the relevant categories of infection risk [15, 16]. The risk of exposure to the pathogen in contaminated ground beef was estimated in both case-control studies by the variable 'eating pink hamburger'. Although participants in the case-control studies were asked about exposure to hamburgers, this variable was not associated with a significantly elevated risk of infection. While consumption of pink hamburger that is not contaminated with STEC O157 will not cause infection, contaminated undercooked ground beef may not appear pink to the consumer [21]. Consequently, the PAF estimated by case-control studies probably underestimates the proportion of cases associated with hamburger consumption [15].

Other factors limiting the number of infections that were attributed to sources in this study were the fairly low estimated risks in the case-control studies associated with common exposures as well as the low prevalence of exposure in cases for higher risk exposures. For example, only 38% of cases were exposed to sources associated with estimated odds ratios exceeding 2·0 in the 1996 study [7]; and in 1999 this proportion was only 18% [8]. These low cumulative attribution proportions suggest that not all relevant

exposure sources were found to be statistically significant in the multivariable model or that the questionnaires did not adequately capture the relevant categories of risk, or both. In addition, not all cases attributed to exposures in the 1996 case-control study were attributed in the blended estimates because these exposure variables, e.g. use of immunosuppressive medication, reflect susceptibility to infection rather than a source. Susceptibility to infection (as estimated by exposure to immunosuppressive medication) is known to modify the risk of foodborne infection following exposure to contaminated sources; thus, there is evidence of biological interaction between the variables included in the final multivariable model. Consequently, we have evidence that the assumption of no interaction between exposures (i.e. contaminated sources of infection) and adjustment factors (e.g. susceptibility) is not valid for the 1996 study, resulting in additional uncertainty regarding the estimated attributable numbers for sporadic infections [9].

Our finding of marked variability in exposure sources between the two years may reflect some or all of the following: true differences, variations in study design, and variations in data resulting from small sample sizes. The impact of small sample sizes is highlighted by the considerable year-to-year variability in the source attribution estimates resulting from large individual outbreaks. For example, in 1996, outbreaks resulting from person-to-person transmission increased the proportion of infections attributable to this route nearly fourfold over the proportion observed in the case-control study of sporadic infections. This was also observed in 1999 when a large outbreak associated with contaminated well water tripled the proportion of infections attributed to untreated water. Although only one outbreak associated with hamburgers was reported during each year of this study, the variability in outbreak size between the two study years resulted in an attribution proportion of 2% of outbreak-associated infections resulting from contaminated hamburgers in 1996 compared to 11% in 1999. In contrast, the proportion of sporadic infections attributable to hamburgers (11%) did not change between 1996 and 1999, and these findings are consistent with the results of ground beef testing by the Food Safety Inspection Service that showed no significant change in the contamination rate of ground beef during this time [22]. It is likely that the contaminated source of exposure influences the number of resulting infections. For example, a contaminated recreational water body is likely to cause more infections than a single undercooked hamburger; using the number of illnesses associated with each outbreak incorporates a source-specific weight in blended source attribution estimates. The stability of this weight can be improved by increasing the sample size and reducing the impact of year-to-year variability in outbreak size on the estimate. If it can be assumed that the sources of infection are similar over a defined time period, blended source attribution estimates may be improved by combining the results of surveillance studies across multiple years. However, these study designs would have to consider whether regulatory changes or consumer habits changed during the study.

Estimates of source attribution can aid in determining appropriate public health interventions and measuring progress toward food safety goals. Data from outbreak surveillance provides the best opportunity to directly attribute illnesses to sources, but most infections are classified as sporadic, thus limiting the generalizability of attribution estimates derived from outbreak data. Including data from both sporadic and outbreak-associated infections can increase the total proportion of infections attributed to an exposure source and improve our ability to generalize estimates to the target population. However, even after combining attribution estimates, a relatively small number of cases were attributed to sources, in part due to methodological challenges. Nonetheless, this study identified several opportunities to improve source attribution using methods that blend information from multiple surveillance systems. For example, models can be developed to estimate differences in ascertainment probabilities in the different surveillance subpopulations so that source attribution estimates may be adjusted before blending, removing the need for the assumption that the exposure distributions are equal in all source subpopulations. Likewise, new approaches may be applied to case-control data that are more suitable for analysing possibly correlated exposures in sporadic infections; and information from outbreak investigations may be used to estimate causation probabilities associated with exposure categories used in case-control studies. These approaches may help to reduce the proportion of unattributed infections, provide estimates of uncertainty for relevant time periods of interest, and contribute to improved source attribution estimation.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None.

REFERENCES

1. **Scallan E, et al.** Foodborne illness acquired in the United States – major pathogens. *Emerging Infectious Diseases* 2011; **17**: 7–15.

2. **Centers for Disease Control and Prevention (CDC).** Foodborne outbreak online database. Atlanta, Georgia: U.S. Department of Health and Human Services, Center for Disease Control and Prevention (http://wwwn.cdc.gov/foodborneoutbreaks). Accessed 23 August 2011.

3. **Alelis KA, et al.** Outbreak of Shiga toxin-producing *Escherichia coli* O157 infection associated with a day camp petting zoo – Pinellas County, Florida, May–June 2007. *Morbidity and Mortality Weekly Report* 2009; **58**: 426–428.

4. **Yoder J, et al.** Surveillance for waterborne disease and outbreaks associated with drinking water and water not intended for drinking – United States, 2005–2006. *Morbidity and Mortality Weekly Report* 2008; **57**: 39–62.

5. **CDC.** Foodbome Diseases Active Surveillance Network (FoodNet): FoodNet surveillance report for 2008 (Final Report). Atlanta, Georgia: U.S. Department of Health and Human Services, CDC. 2010 (http://www.cdc.gov/foodnet/AR2008_508Compliant_v3_FINAL.pdf). Accessed 23 August 2011.

6. **FoodNet.** FoodNet surveillance report for 2009 (Final Report) (http://www.cdc.gov/foodnet/PDFs/FoodNetAR2009_FINAL.pdf). Accessed 22 November 2011.

7. **Kassenborg HD, et al.** Farm visits and undercooked hamburgers as major risk factors for sporadic *Escherichia coli* O157:H7 infection: data from a case-control study in 5 FoodNet sites. *Clinical Infectious Diseases* 2004; **38**: S271–278.

8. **Voetsch AC, et al.** Risk factors for sporadic Shiga toxin-producing *Escherichia coli* O157 infections in FoodNet sites, 1999–2000. *Epidemiology and Infection* 2007; **135**: 993–1000.

9. **Benichou J.** A review of adjusted estimators of attributable risk. *Statistical Methods in Medical Research* 2001; **10**: 195–216.

10. **Bruzzi P, et al.** Estimating the population attributable risk for multiple risk factors using case-control data. *American Journal of Epidemiology* 1985; **122**: 904–914.

11. **Greenland S, Robins JM.** Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology* 1988; **128**: 1185–1197.

12. **Scallan E.** Activities, achievements, and lessons learned during the first 10 years of the Foodborne Diseases Active Surveillance Network: 1996–2005. *Clinical Infectious Diseases* 2007; **44**: 718–725.

13. **Murphree R, et al.** Characteristics of foodborne outbreak investigations conducted by Foodborne Diseases Active Surveillance Network (FoodNet) sites, 2003–2008. *Clinical Infectious Diseases* 2012; **54**: S498–S503.

14. **Knol MJ, et al.** What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *American Journal of Epidemiology* 2008; **168**: 1073–1081.

15. **Levine B.** What does the population attributable fraction mean? *Preventing Chronic Disease* (www.cdc.gov/pcd/issues/2007/jan/pdf/06_0091.pdf). Published online 15 December 2006. Accessed 28 June 2010.

16. **Greenland S.** Attributable fractions: bias from broad definition of exposure. *Epidemiology* 2001; **12**: 518–520.

17. **Rockhill B, Newman B, Weinberg C.** Use and misuse of population attributable fractions. *American Journal of Public Health* 1998; **88**: 15–19.

18. **Steenland K, Armstrong B.** An overview of methods for calculating the burden of disease due to specific risk factors. *Epidemiology* 2006; **17**: 1–8.

19. **Mølbak K, Neimann J.** Risk factors for sporadic infection with *Salmonella Enteritidis*, Denmark, 1997–1999. *American Journal of Epidemiology* 2002; **156**: 654–661.

20. **Rounds JM, et al.** Dynamics of *Escherichia coli* O157:H7 outbreak detection and investigation, Minnesota 2000–2008. *Epidemiology and Infection* 2012; **140**: 1430–1438.

21. **Food Safety Inspection Service.** Meat preparation: color of cooked ground beef as it relates to doneness (http://www.fsis.usda.gov/factsheets/Color_of_Cooked_Ground_Beef/index.asp#1). Accessed 18 October 2011.

22. **Naugle AL, et al.** Food Safety and Inspection Service regulatory testing program for *Escherichia coli* O157:H7 in raw ground beef. *Journal of Food Protection* 2005; **68**: 462–468.