DISCUSSION PAPER

April 2006, updated June 2008 ■ RFF DP 06-17-REV

Attributing Foodborne Illnesses to Their Food Sources

Using Large Expert Panels to Capture Variability in Expert Judgment

Sandra Hoffmann, Paul Fischbeck, Alan Krupnick, and Michael McWilliams

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

RESOURCES FOR THE FUTURE

Attributing Foodborne Illnesses to Their Food Sources: Using Large Expert Panels to Capture Variability in Expert Judgment

Sandra Hoffmann, Paul Fischbeck, Alan Krupnick, and Michael McWilliams

Abstract

Decision analysts are frequently called on to help inform decisionmakers in situations where there is considerable uncertainty. In such situations, expert elicitation of parameter values is frequently used to supplement more conventional research. This paper develops a formal protocol for expert elicitation with large, heterogeneous expert panels. We use formal survey methods to take advantage of variation in individual expert uncertainty and heterogeneity among experts as a means of quantifying and comparing sources of uncertainty about parameters of interest. We illustrate use of this protocol with an expert elicitation on the distribution of U.S. foodborne illness from each of 11 major foodborne pathogens to the consumption of one of 11 categories of food. Results show how multiple measures of uncertainty, made feasible by use of a large panel of experts, can help identify which of several types of risk management actions may be most appropriate.

Key Words: food safety, expert elicitation, risk analysis, food attribution, foodborne pathogen

Contents

Attributing Foodborne Illnesses to Their Food Sources: Using Large Expert Panels to Capture Variability in Expert Judgment

Sandra Hoffmann, Paul Fischbeck, Alan Krupnick, and Michael McWilliams*

1. Introduction

Risk management is ideally guided by good analysis based on sound data. But in many cases, data are lacking or known to be inaccurate and/or biased. In such situations, expert elicitation of parameter values is frequently used to supplement data-based analysis (Cooke 1990; Morgan and Henrion 1990). If the gap to be filled is narrowly defined in terms of discipline or expertise, expert elicitations can safely rely on a small number of experts, particularly if well-performing experts can be identified (Cooke and Goossens 2004). But in many cases, the information needed for risk management must draw on a broad range of disciplines, professional backgrounds, and experience. In these cases, a single expert or even a small group of experts may not be able provide the needed information. Where integration of information is required, it may not even suffice to have a series of expert elicitations that are more narrowly focused. Instead, the elicitation may require a panel of heterogeneous experts.

Van der Fels-Klerx et al. (2002) present a protocol for a formal expert elicitation process to quantify information on continuous variables from a heterogeneous panel of experts. They focus on developing a process to elicit and aggregate expert responses into a single probability density function (PDF) that is as reliable and accurate as possible.

This paper develops a formal protocol for expert elicitation with a heterogeneous expert panel that makes use of individual expert uncertainty and heterogeneity among experts as a means of quantifying and comparing sources of uncertainty about parameters of policy interest. We illustrate use of this protocol with an expert elicitation on the distribution of U.S. foodborne illness across foods. We focus on illnesses caused by 1 of 11 leading foodborne pathogens.

* Sandra Hoffmann is a fellow, Alan Krupnick a senior fellow, and Michael McWilliams a research assistant at Resources for the Future. Paul Fischbeck is associate professor in engineering and public policy at Carnegie Mellon University. This research was funded in part by grants from the Robert Woods Johnson Foundation and the U.S. Food and Drug Administration. The authors wish to acknowledge the contributions of Glenn Morris and Michael Taylor of the University of Maryland School of Public Health, Michael Batz and Juha Siikamäki of Resources for the Future, Michael Roberts of the USDA Economic Research Service, and Jody Tick and Dianne Sherman, former research assistants at Resources for the Future.

Section 2 provides background on this application. In Section 3, the formal protocol is described, together with its application in the food safety context. Results and implications are presented and discussed in Sections 4 and 5, respectively.

2. Background

U.S. food safety policy is organized around specific pathogens on specific foods. But U.S. food safety policymakers do not have good information on how pathogens are distributed in the food supply or on the relationship between food consumption and illnesses caused by specific pathogens. Current U.S. estimates of the incidence of foodborne illness by pathogen rely on outbreak data and active and passive surveillance data collected by the Centers for Disease Control (CDC) and state epidemiologists (CDC 2006). The only existing comprehensive attribution of foodborne pathogen illness to foods in the United States is based on CDC outbreak data (DeWaal and Barlow 2002). Foodborne-illness surveillance systems do not regularly collect data on food consumption associated with illness, and outbreak data do not provide a reliable picture of the distribution of illness by food-pathogen pairings. Moreover, the quality and consistency of outbreak data vary by state, because state governments direct these investigations.

At best, outbreak data give a lower bound on the number of cases of foodborne illness by food-pathogen combination, measured with recognized but unquantified error. Sporadic cases of foodborne illness are not included in outbreak data. Error enters for many reasons. State departments of public health collect outbreak data which results in variation in data collection. Cases of rare or difficult-to-identify pathogens are often underrepresented. Cases from large outbreaks that are associated with restaurants or cases associated with unusual food vectors tend to draw attention and are believed to be over-represented.

Whether or not outbreak data represent a lower bound to the number of cases, there is little reason to believe that outbreak data provide a lower bound on estimates of the *percent* of U.S. foodborne illness cases caused a pathogen-food combination. The only consistent attribution of foodborne illness caused by specific pathogens to food consumption uses outbreak data. But sporadic cases are estimated to account for more cases of foodborne illness than outbreak cases by an order of magnitude or more (Mead et al. 1999). No such good data exists on how sporadic cases are distributed across foods and it is unlikely that that data will become available soon in the U.S. system.

A fundamental problem is that direct evidence of the association between foodborne illness and food consumption in sporadic cases has often been destroyed or consumed by the time the illness is reported and is very seldom collected even when the evidence is available. Furthermore, dietary recall is notoriously inaccurate (Carter et al. 1981; Bernard et al. 1984). In addition, the risk-generation process that results in outbreaks, which by definition are rare events, could be very different from that that results in sporadic case. As a result, the distribution of sporadic and outbreak cases across foods could differ substantially. It is perfectly possible that for any food-pathogen combination, the actual percentage of total foodborne cases of illness caused by a particular pathogen is either higher or lower than the percentage of outbreak cases.

Food safety experts from a wide variety of backgrounds have substantive knowledge that could inform a sound judgment about the percentage of a pathogen causing disease in various foods. Knowledge of pathology, epidemiology, microbial ecology, veterinary medicine, food production, processing and storage technology and practices, food marketing and handling practices, food consumption patterns, and human medicine are all relevant to an understanding of the likely association of illness with consumption of particular foods. Despite the wide range of data that may be relevant to informing a judgment about this relationship, there is no directly comparable task that can be used to test experts' performance. In such a situation, expert elicitation cannot reliably turn to calibration of responses based on a similar known estimation task (Cooke 1991). Risk managers would like to have a sound estimate of this relationship, but they are equally interested in knowing where existing data collection efforts are weak and where to focus future data collection efforts.

3. Elicitation Protocol

The protocol developed here makes use of the kind of heterogeneity of relevant backgrounds seen in the food safety context. This kind of situation, where multidisciplinary expertise is relevant to the risk assessment, is common to many practical decision-analysis problems. We supplement Van der Fels-Klerx et al. (2002) by focusing on uncertainty and using formal survey methods. These innovations are motivated by the need for an expert elicitation protocol that allows systematic examination of uncertainty where a broad range of expertise is relevant and no good means exists for directly evaluating the quality of data or expert judgment. The protocol has seven basic steps that are a modification of standard expert elicitation protocols (Clemen 1996; Cooke and Goossens unpublished manuscript):

1. Determine the size of the expert panel.
2. Choose the model of analysis and/or aggregation.

3. Choose the mode of elicitation.

4. Develop the elicitation survey instrument:

- identify relationships to be elicited and categorize hazards;
- decide how to manage experts' information set;
- decide how to motivate expert judgments and address cognitive biases;
- decide on relevant independent control variables; and
- pretest the elicitation instrument.

5. Identify the expert pool.

6. Administer the elicitation survey.

7. Analyze the survey results.

3.1. Determining the Size of the Expert Panel

The size of the panel depends on 1) the type of problem for which expert judgment is being elicited; 2) the nature and degree of uncertainty about the problem; and 3) the range of relevant expertise needed to assess the problem. The narrower the task and the more defined the nature and extent of uncertainty, the narrower is the required the range of expertise. In such circumstances a small expert panel is adequate. Expert elicitations usually rely on relatively small expert panels, typically fewer than a dozen respondents (Merrick et al. 2005; Winkler and Clemen 2004). But the broader and more complex the problem, or the greater the degree of uncertainty about the task, the larger the panel will likely need to be. Van der Fels-Klerx et al. (2002), for example, use a heterogeneous expert panel of 15 to 22 experts to identify the factors relevant to estimating the impacts of bovine respiratory disease on dairy heifer production. Keeney and von Winterfeldt (1991) use a panel of 40 experts to assess the probabilities of failure associated with alternative sites for high-level radioactive waste storage. We believe our problem is best characterized as a broader and more complex problem and therefore requires a larger panel.

3.2. Mode of Analysis and Aggregation

Much of the literature on expert elicitation focuses on how to aggregate expert judgment to achieve a reliable consensus on parameters or distributions of interest (Cooke and Goossens 2004). Broadly speaking, two approaches have been taken to aggregation: behavioral and mathematical. Delphi processes are a widely used behavioral approach to aggregation that use highly structured, iterative group processes to reach consensus among experts. Van der Fels-Klerx et al. (2002) use this approach in aggregating heterogeneous expert panels. Several alternative means of aggregating expert judgment through weighting expert judgments have been developed, including the use of a seed variable as a means of measuring experts' performance (Cooke and Goossens 2004). Depending on the purpose of the elicitation, aggregation may not be appropriate. In many expert elicitations, the range of expert opinion may be at least as valuable as the aggregate assessment (Keith 1996). In cases with a high level of uncertainty, such as climate change, Lempert et al. (2004) have argued that it may actually be misleading to attempt to construct even a consensus probability distribution. The breadth of our problem implies that questions that can rank a wide range of experts as to the quality of their judgments will be impossible to identify.

3.3. Mode of Elicitation

There is an interaction between the desired type of analysis or aggregation method and the elicitation approach. The protocol developed in this study is intended to be used in situations of high uncertainty, where there are nonetheless a large number of experts with relevant knowledge. In this setting, a formal elicitation survey and standard statistical analysis of the survey responses can be used to preserve and formally analyze the range of expert opinion.

3.4. Development and Administration of an Elicitation Instrument

In the development and administration of an elicitation survey instrument, all of the basic questions encountered in other forms of expert elicitation must still be addressed (Clemen 1996). These include: identifying the relationships on which judgments are sought, classifying hazards, identifying the expert pool, controlling cognitive biases, deciding when and how to influence the information set experts bring to the elicitation task, recruiting experts, and motivating the assessment task. In addition, standard concerns that arise in administration of any survey, such as avoiding sample bias and ensuring an adequate response rate, must be addressed. An acknowledged hazard in expert elicitations is that experts may interpret questions differently (Morgan and Henrion 1990). This problem is not unique to expert elicitation: if a formal questionnaire is used, survey methodology addresses this problem through pretesting and revision of the survey instrument (Rea and Parker 1997).

In our foodborne illness application, a survey instrument was developed that consisted of a demographic questionnaire on educational background and professional experience, a self-evaluation of food-specific expertise, a self-evaluation of pathogen-specific expertise, a hypothetical food attribution worksheet, and separate food attribution worksheets for each pathogen. The survey instrument, together with a cover letter explaining the project and instructions for the survey, was pretested on several food safety professionals in the Washington, DC, area.

• *Identification of Relationships to Be Elicited and Hazard Categorization.* Morgan et al. (2000) call for mutually exclusive and administratively relevant categorization schemes in comparative risk assessments and rankings. As U.S. food safety policymakers strive to take a more risk-based approach to regulatory decisionmaking, one missing piece of information needed for priority setting is a solid understanding of the distribution across foods of illnesses by causal pathogen. In the example elicitation, experts were asked to attribute cases of illness caused by each of 11 major foodborne pathogens to consumption of food from one of 11 categories that represent broad classes of food at the point of consumption (Figure 1). Categorizing foods at the point of consumption provides categories that are meaningful to the patients and physicians who ultimately report illnesses to CDC. This approach also provides a direct association between pathogen illness and food at the point of exposure, which facilitates risk analysis and is consistent with the focus of regulatory programs. We use food categories developed by the Center for Science in the Public Interest (CSPI) in its annual attribution of CDC outbreak data. This allows us to compare expert judgments with existing outbreak attributions. The CSPI categorization scheme is quite generic and was viewed as generally unobjectionable in the survey pretests by food safety experts.[1]

The pathogens included in the study are the nine pathogens CDC follows in FoodNet, its primary foodborne illness surveillance program—*Campylobacter spp.* (Campylobacter), *Cryptosporidium parvum* (Cryptosporidium), *Cyclospora cayetanensis* (Cyclospora), *Escherichia coli* O157:H7 (E. coli *O157:H7)*, *Listeria monocytogenes* (Listeria), *Salmonella nontyphoidal* (Salmonella), *Shigella*, *Vibrio spp.* (Vibrio), and *Yersinia enterocolitica*

---

[1] CSPI also includes a category for other/multi-ingredient foods. This category overlaps with other food categories and is therefore not mutually exclusive. Survey pretests indicated that expert respondents felt able to attribute illnesses based on food ingredients, given their knowledge of microbiology, food-handling practices, and food consumption patterns. For example, outbreak data could report potato salad as a multi-ingredient food, but food safety experts could make the judgment that there was a much higher probability that the problem was the raw egg, not the boiled potato. As a result, we do not include a multi-ingredient category. We include the category "other" in data collection but drop it in the final analysis, for which the remaining attributions are scaled up to add to 100 percent.

(Yersinia)—plus *Toxoplasma gondii* (Toxoplasma) and Norwalk-like viruses (Norwalk) (CDC 2003). The inclusion of pathogens in the FoodNet data collection effort is an indication of their public health importance. In addition, Toxoplasma is estimated to account for the third-highest number of deaths from foodborne pathogens in the United States. Norwalk is estimated to account for the highest number of cases (Mead 1999).

• *Information Set.* Expert judgment is inherently conditional on an information set. The extent to which expert elicitations attempt to influence this information set is largely determined by the purpose of the elicitation and the underlying state of relevant scientific knowledge. When parameter values are sought for a reasonably well-understood process, elicitations have provided extensive information packages and briefings on which experts are instructed to base their judgments (Cooke and Goossens 2004). In other cases, like the food attribution used as an example in this study, the need is to understand the information that experts bring to the task (Lempert et al. 2004).

• *Cognitive Biases and Elicitation of Uncertainty.* Expert elicitations need to address common cognitive biases that could skew elicited responses. Respondents' judgments tend to be anchoredon extreme values or on information first presented or elicited (Tversky and Kahneman 1974). In addition, they are often overconfident of their knowledge and provide confidence intervals around an elicited estimate that are overly narrow. Several "debiasing" techniques (e.g., assessing endpoints first, thinking about extreme events, and discussing the nature of the bias) have been used to help widen confidence intervals (Epley and Gilovich 2001; George, Duffy, and Ahuja 2000; Morgan and Henrion 1990). These techniques have not been applied to the kind of task asked of experts in this paper—that is, assessing a vector of proportions that sum to 1.0 and providing a confidence interval around each proportion. This is a much more constrained task than those typically given in debiasing exercises.

Anchoring of some type is impossible to avoid. The survey design in our example elicitation focuses the respondent's attention first on the "best" estimate and then on the extreme values of their confidence intervals (Figure 2). Because the best estimates have to sum to 1.0, eliciting confidence intervals, first followed by best estimates, would be cognitively difficult and could have increased the refusal rate significantly. Elicited confidence intervals, or ranges, are used to make within-expert and between-expert comparisons. Providing that the respondents are similarly susceptible, these comparisons should be relatively unaffected any bias.

• *Independent Variables as Indicators of Background or Expertise*. Like other decisions, expert judgments are based on respondents' mental models of the decision problem (Morgan et al. 2001). Mental models are small-scale, subjective representations of external reality (Craig 1943). They include information sets, evaluative criteria regarding the relevancy of information, and understandings of causal associations (Gentner and Stevens 1983). The learning and socialization inherent in education and professional experience influence experts' mental models of risk-generation processes (Morgan et al. 2001). Past expert elicitations have found probability assessments to vary by field of expertise (Banke and Jenkins-Smith 1993; RFF 2006). Survey questions can be used to evaluate how expert judgment varies systematically by factors that could affect respondents' mental models of the risk generation process. In this food attribution example, we collected data on each respondent's field of study, professional work setting, experience, and self-evaluated expertise.[2]

3.5. Identification of Expert Pool

A systematic approach to expert selection is needed to avoid bias. Cooke and Goossens (2004) recommend that experts be chosen on the basis of reputation, experience, and publications. Morgan et al. (2002) recommend that a panel of experts be as balanced as practicable in terms of affiliation, training, and subject matter. Balance across background is particularly important when selecting a large, heterogeneous expert panel to characterize uncertainty about parameters of interest.

In our food safety example, potential respondents were chosen on the basis of extensive professional experience in food safety and recognized expertise in relevant areas of food safety. Care was taken to include people representing a wide range of scientific background and professional experience. Relying on an individual's publication record as a primary criterion for level of expertise would have eliminated state epidemiologists, who run the public health surveillance programs and have considerable expertise about the quality of existing estimates. Twelve leading experts in food safety science and management were asked to review the initial and revised lists of potential respondents. One major result of this review was the inclusion of state epidemiologists with food safety responsibilities. (State epidemiology offices have primary

---

[2] Respondents were asked to rank their own expertise for each pathogen on a 5-point Likert scale, with 1 being low and 5 being high. A low level of expertise was defined as "no direct experience, anecdotal knowledge only." A medium level of expertise was defined as "some direct experience and wide reading." A high level of expertise was defined as "primary focus of my professional work." Respondents were asked to fill out a similar questionnaire for each food category. Use of self-evaluated expertise to eliminate experts or as a weight in expert elicitation has generally been rejected (Cooke 1991). However, we used it as a means of examining factors influencing the degree of an individual respondent's uncertainty about his or her own best attribution estimate.

responsibility for investigation of foodborne illness outbreaks.) A pool of 100 possible respondents was ultimately identified through this iterative review process. We were unable to obtain usable contact information for 11 of the possible 100 respondents.

3.6. Administration of the Elicitation

Choice of how to administer a survey entails a trade-off between cost and response rate. Typically surveys administered in person have the highest response rates. In-person administration also allows for probability training and assessment techniques developed elsewhere in the expert elicitation literature (Morgan and Henrion 1990; Cooke 1991). But this approach is costly. Mail surveys are relatively inexpensive and have the advantage of not introducing interviewer bias, but response rates are typically low. An intermediate approach is to use a mail survey with phone solicitation and follow-up.

We used a mail survey with phone solicitation and follow up in administering our food safety expert elicitation survey. After the purpose of the survey and the nature of the elicitation task was explained, 27 respondents on the list of 89 potential respondents declined to participate, primarily because they felt they lacked the necessary expertise. Of the 62 respondents who agreed to participate and were sent surveys, 44 (67 percent) returned completed surveys. This response rate is a good even for mail surveys of subpopulations receiving surveys of high salience, such as government surveys of business establishments (Groves et al. 2004; Lu 2004).

Because of the breadth of the attribution task, respondents were instructed that they did not have to respond on particular pathogens for which they felt they did not have adequate expertise. Responses by pathogen ranged from a low of 34 to a high of 40. On average there were 38 responses per pathogen. In general, the basic quality of responses was good: of the 432 pathogen level responses, 94 percent met adding-up constraints. Only 5 of the 4,488 range responses were illogical. Together this indicates that respondents found the expert elicitation survey no more difficult to complete than other mail surveys and were able to provide responses that are consistent with probability theory.

4. Analysis and Results

4.1. Respondent Pool

One of the basic design goals was to survey a relatively large group of experts from a wide range of relevant backgrounds. The pool of 100 potential respondents represented a broad range of workplaces. Forty-two were working in government, with 15 in state government (primarily state epidemiologists); 24 in federal government; and 3 in other government positions. Among those in federal government, nine work at the Food and Drug Administration (FDA), seven at the U.S. Department of Agriculture (USDA), four at CDC, and four in other offices and agencies. Thirteen respondents were working in industry and related private-sector employment. Forty-three were employed in academic institutions. Two were working in other settings.

The final respondent pool was reasonably representative of the original list of potential respondents. Three respondents reported having significant work experience in multiple institutional settings. The remainder were evenly distributed among government, academia, and industry. The response rates relative to the list of 100 potential respondents (which includes 11 people who could not be contacted) were 60 percent for government, 62 percent for industry, 35 percent for academia, and 100 percent for other. Twenty-four (54 percent) of the respondents listed government as the principal setting in which they had worked over their careers. Eight respondents (18 percent) currently work in state government and most of them were state epidemiologists. Respondents who currently work for the federal government were spread over the primary agencies with food safety responsibility: six at FDA, five at USDA, and three at CDC. Fourteen respondents (32 percent) said that academia had been their primary place of employment over their careers. Three respondents (7 percent) listed industry as their primary place of employment, and two listed "other."

The respondents had significant professional experience and training in food safety. Of the 44 experts who completed the survey, 63 percent had 20 years or more of professional experience with this issue; 77 percent had at least 12 or more years of experience. On average, respondents had 21.5 years of experience working on food safety. Respondents received their highest degrees in a broad range of fields relevant to food safety: medicine (25 percent), food science (18 percent), public health (17 percent), microbiology (17 percent), and veterinary medicine (11 percent).[3] Ninety-four percent of the sample had PhDs, MDs, or DVMs. Of these

[3] Respondents were asked their field of work in an open-ended question. Responses were then coded to 6 fields of work as follows: Microbiology Group (bacteriology, food science microbiology, microbiology, area pathology), Public Health Group (public health, public health epidemiology, epidemiology), Veterinary Medicine Group (veterinary medicine, veterinary medicine parasitology), and Other (ecology, math, nutrition, quality management). Food science and medicine were retained as independent groups.

there were 23 PhDs, 13 MDs, 3 DVMs, and 1 DVM and Ph.D. The remaining respondents had other relevant graduate degrees.[4]

4.2. Self-Evaluated Expertise

All respondents provided self-evaluations of expertise on foods, and all but one provided a self-evaluation of expertise on pathogens. Experts self-evaluated as having the highest level of expertise on meats, particularly beef and poultry (Table 1). Produce also ranked high. Beverages, breads and bakery, and game ranked lowest. Historically, meats and produce have been seen as posing greater public health risks than beverages and breads and bakery (foods whose hazards are naturally lower or more readily controlled) and game (which is not widely consumed). Respondents also believed they understood some pathogens better than others. High on this list were *E. coli* O157:H7, *Listeria*, *Salmonella*, and *Toxoplasma*; *Cryptosporidium* and *Cyclospora* have low mean expertise scores.

4.3. Graphical Analysis of Uncertainty Measures

We use three measures to characterize uncertainty about the attribution of pathogen-related illnesses to foods: 1) the extent of agreement between expert and outbreak-based attributions; 2) the degree of variability among experts' best estimates; and 3) individual uncertainty. Given the quality of data and uncertainty about the relationship between foodborne illnesses cause by particular pathogens and food consumption, it is not possible to say whether outbreak-based attributions or attributions from this elicitation are correct. Yet it seems plausible that there may be less uncertainty about attributions where experts agree strongly with outbreak-based attributions.

Figure 1 provides a graphical overview of the first two measures of uncertainty—agreement with outbreak data and variability among experts' best estimates. Four major patterns emerge. First, more difference occurs among food-pathogen combinations than across pathogens or for particular foods. Second, there is substantial agreement between expert and outbreak-based

[4] Only a few reasonably strong correlations emerge among respondent-related variables (Pearson product-moment correlations that exceed 0.5 in absolute value). Not surprisingly, having a career primarily in government is negatively correlated with one in academia (–.73). But having a career in any of the other sectors has a negative but low correlation with having a career in other sectors. Having a PhD is positively correlated with working in microbiology (.52) and negatively correlated with working in medicine (–.58) or having an MD (–.67). Respondents with DVMs predominantly work in veterinary medicine (.91) and those with MDs in human medicine (.87).

attributions about which food-pathogen combination are not problems (i.e., when the attribution rate is low). In general, a food-pathogen combination attributed with close to zero percent of cases has interquartile ranges close to zero. This observation is examined analytically below. Third, attributions do vary by pathogens.

At one extreme are two pathogens, *Vibrio* and *Cyclospora*, for which there is also close agreement between expert and outbreak-based attribution on foods attributed with a high percentage of cases. These pathogens have single, well-understood transmission pathways: *Vibrio* is found almost exclusively in seafood, and *Cyclospora*, almost purely waterborne, contaminates raw produce.[5] At the other extreme are three pathogens for which there is substantial disagreement between expert and outbreak-based attribution: *Campylobacter*, *Toxoplasma*, and *Cryptosporidium*. These three pathogens differ in the degree of variation among experts.

The remaining pathogens have a reasonable degree of agreement between expert and outbreak-based attribution but can be grouped by level of expert variability. *Shigella* and Norwalk-like viruses have fairly large inter-quartile ranges for foods attributed with more than 10 percent of caseload by experts. The remaining four pathogens—*E. Coli* O157:H7, *Listeria*, *Salmonela*, and *Yersinia*—have more moderate inter-quartile ranges. Finally, most foods have only one pathogen for which either the difference between expert and outbreak-based attribution or the inter-quartile range exceeds 10 percentage points. Three foods have two or more pathogens where this holds.

4.4. Statistical Analysis of Uncertainty Measures

• *Agreement between Expert and Outbreak-Based Attribution.* There is substantial agreement between experts and the outbreak-based data about cases in which a food-pathogen combination is not a problem. Both experts and outbreak-based estimates attribute zero percent of cases to 47 foods (out of 121 possible combinations). For 22 of the food-pathogen combinations, the mean expert best estimate is greater than zero but less than five percent when the outbreak-based attribution is zero. In total, then, experts and outbreak-based estimates agree that more than half (69) of the food-pathogen combinations account for less than 5 percent of the cases of foodborne illness caused by a particular pathogen.

[5] Personal communication with Dr. Glenn Morris, University of Maryland School of Public Health.

There are a substantial number of food-pathogen combinations, however, that account for more than a *de minimis* proportion of cases and for which experts do not agree with the outbreak-based attribution. Table 2 presents descriptive statistics and tests of means for 43 food-pathogen combinations for which one of the two estimation methods attribute at least 5 percent of pathogen-related cases. For 28 of the 43, more than half of the 5th to 95th percentile ranges do not include the outbreak estimate. Over all possible pairings (not shown), 50 ranges (5–95th percentiles) do not include the outbreak-based estimates.

Another way to look at the comparison of expert and outbreak estimates is to ask whether the ranking of foods for each pathogen is the same. As suggested by the box-and-whisker plots, for several pathogens the ranking of foods in terms of the percentage of illnesses differs substantially for expert and outbreak-based attributions. The most dramatic cases are *Campylobacter*, *Cryptosporidium*, and *Toxoplasma*. Outbreak-based attribution characterizes *Campylobacter* as a problem in produce and diary; experts see it as overwhelmingly associated with poultry. Outbreak-based attribution associates *Cryptosporidium* cases chiefly with beverages and secondarily with produce; experts would reverse this ranking and spread cases among a much wider range of foods than implied by outbreak data. Outbreak data associates 100 percent of cases of *Toxoplasma* with game, but because this reflects a single outbreak, it is unsurprising that experts believe cases are more widely distributed. For several pathogens, including Norwalk and *Listeria*, *Salmonella*, *Yersinia* and *Shigella,* there are less marked differences. And for others—in particular, *Vibrio, Cyclospora,* and *E. coli* O157:H7—there is strong agreement between expert judgment and outbreak-based attributions.

We use regression analysis to explain differences between an individual expert's best estimates and the outbreak data estimates. As just discussed, difference may vary by food, pathogen, or food-pathogen pairing. They may also vary by expert characteristics such as career affiliation, final academic degree, field of expertise, years of experience, or self-assessment of expertise. Finally, we hypothesize that additional outbreaks associated with a specific pathogen add information and decrease the difference between expert and outbreak-based attributions.

Table 3 presents results of Tobit regressions of expert best estimates minus outbreak-based attributions testing these hypotheses. We use Tobit regression to accommodate the censoring of the dependent variable at 0 and 100. Stepwise examination of possible hypotheses was used to identify a final model.

Self-assessed food expertise was significant in explaining the difference between attribution estimates (at the 1 percent level). A one-point increase in mean self-reported food expertise is associated with a 1.35-percentage-point increase in the absolute difference between expert and outbreak-based attribution. But this significance disappears when food dummies are added to the equation. Although there is not significant correlation between food expertise and individual dummy variables, sign reversals and change in significance of coefficients for foods suggest multicollinearity is affecting results. Thus it is not possible to distinguish between the influence of expertise and food dummies. In addition, the number of outbreaks is significant in all models at a 1 percent level. An additional outbreak results in a 0.05 to 0.08 percentage point decrease in the difference between attribution estimates, consistent with the idea that more outbreaks leads to greater confidence among the experts that the outbreak data accurately represent reality.

The difference between attribution estimates varies systematically by food and pathogen. Columns I and II of Table 3 report the influence of pathogen identity and food type on this difference without including other variables. The omitted dummy variables are *Campylobacter* and produce. The mean difference for produce is larger than for any other food, ranging from 9 percentage points larger than poultry to 35 percentage points for eggs. Table 3 shows that on average, the difference between attribution estimates is 45 percentage points smaller for *Cyclospora* than for *Campylobacter*. Hoffmann et al. 2006 reports results in which alternative pathogen and food dummies are omitted. When *Cyclospora* is omitted as a dummy variable, the difference between attribution estimates for *Cyclospora* is statistically different from all other pathogens except *Vibrio* at the 1 percent level. The difference between expert and outbreak-based attributions is 12 to 35 percentage points smaller for eggs than for any other food. Likelihood ratio tests comparing the full mixed-effects model with restricted models find that food/pathogen combination dummies jointly and food dummies jointly are significant at the 1 percent level. Pathogen dummies are not jointly significant even at the 10 percent level.

• *Variability across Experts*. Variability among experts' judgments also provides an indication of whether knowledge about a relationship is settled. This is a statistical measure of consensus of expert opinion. The standard deviation of the best estimates ranges from a low of 0 for 13 food-pathogen pairings to a high of 35 percentage points for Norwalk-like viruses in seafood. Figure 3 shows that variation among experts' best estimates is lowest for food-pathogen combinations for which experts have low mean best estimates and those few with very high mean best estimates. This is also reflected in the fact that the median standard deviation is 4.24 and the 75th percentile is 8.47. Table 4 lists the food-pathogen pairings in the top decile of standard deviation of best estimates. These include some of the most studied of food/pathogen combinations, including *Listeria* in dairy products and luncheon meats and Salmonella in poultry, as well as some of the least studied, such as *Toxoplasma* in game, pork, and beef.

Regressions show that the curvature seen in Figure 4 is statistically meaningful (Table 5). All models show the standard deviation of best estimate as being quadratic in mean best estimate. There are a number of reasons for this. One is that as best estimates approach the boundaries of 0 and 100, there is less "room" for disagreement because best estimates are bounded either above or below. In addition, at best estimates of 0, there is a great deal of agreement among experts about which foods are not vectors for particular pathogens. To a lesser extent, there is also significant agreement on those cases where experts think there is a single food that is the predominant vector for a particular pathogen. Beyond this curvature, Table 5 shows that other factors explain the standard deviation of best estimates. The number of outbreaks is significant. Higher mean self-reported pathogen expertise, but not mean self-reported food expertise, is associated with lower variability among experts' best estimates. Columns I and II of Table 5 look at the influence of food and pathogen dummies in isolation, with *Salmonella* and breads omitted. Differences in variability by pathogen and by food seen in Figure 2 do appear to be statistically meaningful. For example, the mean standard deviation for *Salmonella* is not distinguishable from that for *Listeria* or *E. coli* O157:H7, but it is statistically smaller than that for *Toxoplasma* or *Shigella*. Similarly, the mean standard deviation for seafood and game is statistically different from and higher than, that of breads, eggs, or several other foods (Table 5).

• *Individual Uncertainty*. Experts also vary in how certain they are about their own attribution judgments. Individual uncertainty is measured as the difference between the 95 percent and 5 percent confidence bounds that individual experts put around their own best estimates. We call this measure the "range" of the experts' attribution estimates. Because upper and lower confidence bounds cannot be constrained to sum to 100, they cannot be directly compared across individuals or food-pathogen pairings. However, the size of the difference between these bounds does give information on the individual expert's degree of certainty about his or her estimates and can be compared across individuals and food-pathogen pairings.

Figure 4 plots range by pathogen and by food. Again, variability across food-pathogen combinations is greater than by pathogen or food on average. As with best estimate and standard deviation of best estimate, small ranges predominate. Only 29 of the total 121 food-pathogen combinations have ranges of 10 percentage points or more. All of these are for combinations attributed with at least 5 percent of pathogen-related cases by either the expert or the outbreak-based attribution (Table 2). Although there is variation by pathogen in ranges at the food-pathogen level, and in aggregate, measured as mean range by pathogen, no pathogen dominates. In contrast, the mean range for produce is more than double that of the next-highest food, and produce has a range greater than 10 percentage points for 8 of the 11 pathogens and individual uncertainty about breads and eggs is very low. Across all foods and across all pathogens, average range is 6.98 percentage points, ranging from zero to 46.6 percentage points. Figure 5 plots experts' mean range by food-pathogen combination against experts' mean best estimate by food-pathogen combination. As in Figure 4, the plot is distinctly quadratic.

Table 6 reports results for Tobit regressions for range. In contrast to the regressions on the difference between attribution estimates and variability among expert best estimates, both respondent-related variables and food- and pathogen-related variables help explain variation in average range. The first two columns of Table 5 report results for Tobit regressions of range on pathogen dummies, excluding *Toxoplasma,* and on food dummies, excluding produce. The results again support results found earlier on the difference between attribution estimates or variability in best estimates (Figure 6). Mean expert ranges for *Cyclospora, Vibrio,* and *Yersinia* are smaller than for *Toxoplasma.* Mean ranges for several pathogens are indistinguishable from that of *Toxoplasma*, including *Campylobacter* and *Cryptosporidium* as well as *E. coli* O157:H7, *Shigella*, *Listeria*, and *Norwalk*. Regressions confirm that the range for produce is statistically different from and larger than that of other foods.

In contrast to our earlier results, experts' backgrounds and experiences as well as self-reported pathogen expertise help explain variation in individual experts' ranges. Respondents who identify government as their primary career setting have tighter ranges than those whose careers have been primarily in academia, industry, or multiple sectors. Those with significant career experience in multiple sectors have the largest ranges (11.05 percentage points larger than government respondents), followed by those in industry (4.6 percentage points greater than government), followed by academia (2.3 percentage points greater than government). This result is basically invariant to model specification. Highest degree also explains variation in range. Those with master's degrees have the least confidence in their best estimates, and DVMs have the most. On average, the ranges of DVMs are almost 7 percentage points smaller than those of experts with master's degrees. The averages ranges of MDs and PhDs are roughly 3 percentage points smaller than for those with master's degrees. The field of expertise also matters. Relative to microbiology (bacteriology, food science microbiology, microbiology, and area pathology), those who identified their field as "public health" (public health, public health epidemiology, and epidemiology) have ranges that are 4 percentage points larger than the microbiology group. Veterinary medicine had ranges that are 3 percentage points smaller than the microbiology group. In the more complete models, an increase in pathogen expertise, but not food expertise, is associated with a statistically significant decrease in range. When food and pathogen dummies are added, the coefficient of pathogen expertise becomes smaller but is still significant. The addition of random effects introduces significant collinearity and so is not reported.

5. Discussion

The foregoing results show a wide range of variation at the food/pathogen level. Without a large sample size, we could not have tested to see if uncertainty varied systematically by discipline or professional background. We see that variability in best estimates, but not individual uncertainty or difference between expert and outbreak attributes does differ by professional background and discipline. This association is consistent with the mental models literature.

Use of a larger expert panel also allows us to develop multiple measures of uncertainty: difference between expert judgment and existing estimates, individual uncertainty, and variability among experts. These measures capture different aspects of uncertainty that are meaningful to risk management decisions. Figure 7 shows how food/pathogen pairings can be assigned to one of eight cells representing alternative combinations of high and low levels of uncertainty based on these three uncertainty measures. These relationships form a three-dimensional cube, but for ease of interpretation we unstack the cube in Figure 7. In general, there is a strong positive relationship between mean range and standard deviation of best estimates by food pathogen combinations (Pearson correlation coefficient of 0.86). As a result, for this expert elicitation, the cells with low uncertainty and low variability and the cells with high uncertainty and high variability will be more heavily populated than those with low uncertainty and high variability or high uncertainty and low variability.

The eight comparisons of uncertainty measures represented in Figure 7 can be used to inform risk management decisions. Where individual experts' uncertainty is low on average and experts agree with one another in their best estimates, decision makers should feel most confident about the state of knowledge about the elicited value. In our food attribution survey, there are many such food/pathogen combinations, particularly those attributed with few or no cases of illness. Strong agreement with prior estimates, in this case outbreak-based attributions, would add to that confidence. Where expert judgment does not agree with prior judgment, experts share significant information that is not captured in the prior judgment. In our food attribution survey, these are food/pathogen combinations where experts are confident and unified in their disagreement with outbreak-based attribution estimate, such as *Yersinia* on luncheon meat or *Campylobacter* on seafood. In such cases, meta-analysis or focused literature-based risk assessments may be adequate to gain a sound understanding of the association between food consumption and pathogen contamination. In cases where experts are very confident about their best estimates but there is significant variability among experts, risk managers might consider doing more focused expert elicitations. Expert elicitations using group process to uncover why experts disagree and yet individuals are quite certain. In our food attribution expert elicitation, *Cyclospora* in game or *Campylobacter* in eggs are examples of low individual uncertainty, but high variability among experts. Cases where there is significant individual uncertainty, but little variability in responses, even if there is high agreement with outbreak attributions, suggests that additional epidemiological and other primary research might be warranted to reduce the uncertainty. An example is *Cyclospora* on produce in our survey. The quadrant with high uncertainty and high variability more strongly suggests the need for primary research, particularly if many where the attribution estimates are high. Examples from our survey include *Cryptosporidium* or *Shigella* on produce. Results from use of a protocol like the one developed here would provide the basis for value of information analysis that would be useful in deciding whether it is worth spending research money to reduce this uncertainty.

6. Conclusions

This study provides some basic insights into use of formal survey methods to elicit probability beliefs from large heterogeneous expert panels. Fundamentally, we have learned that experts can do probability assessment in a mail survey context. Reasonably high response rates and low level of spoiled surveys suggest that this is an approach to probability assessment that experts will accept and can perform. We also see that statistical analysis of a large heterogeneous sample can help more efficiently set research and policy agendas. A number of areas for further research would improve this approach. More research is needed on how to address cognitive bias in a survey setting. Split sample studies would be helpful in understanding whether survey, in-person, and group process elicitation methods result in different probability assessments. Because sampling frame is important with larger, more heterogeneous panels, research on how to identify the panel would be useful. Such research could draw on the social networks literature to develop optimal strategies for identifying a set of potential expert respondents.

Decision analysts are frequently called on to help inform decisionmakers in situations where there is considerable uncertainty. Sometimes, this uncertainty is limited and reasonably well understood, but the relationship between particular variables or the distribution of particular parameters needs to be better understood. Sometimes—as in the case of climate change or, surprisingly, in the case of managing pathogen-related foodborne illness in the United States—the uncertainty is great enough that a better understanding of the nature of that uncertainty is needed to advance policymaking. When such situations also require understanding a broader range of relationships that touch on a wide range of expertise, the universe of relevant experts will be heterogeneous. In both situations, expert judgment can be a useful supplement to hard data. Other researchers have developed protocols that use heterogeneous expert panels to quantify parameter distributions. This study shows how formal survey methods and statistical analysis can take advantage of the greater heterogeneity of expert background and opinion to gain a better understanding of where further effort is needed to reduce uncertainty.

This paper has focused on the use of a survey protocol to examine uncertainty on a wide but related set of relationships between foodborne pathogens, illness, and food consumption in the U.S. food supply. A next step is examining the value of additional information to reduce the uncertainty identified in this paper. The same survey protocol can also be used to quantify distributions on these parameters values. Hoffmann et al. (2006) use data collected in this same survey to examine implications for food safety policy priorities.

References

Barke, R. P., H. C. Jenkins-Smith. 1993. Politics and Scientific Expertise: Scientists, Risk Perception, and Nuclear Waste Policy, *Risk Analysis.* **13** 425–429.

Bernard, H. R., P. Killworth, D. Kronenfeld, L. Sailer. 1984. The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology.* **13** 495–517.

Carter, R. L., C. O. Sharbaugh, C. A. Stapell. 1981. Reliability and Validity of 24-hour Recall. *J. Amer. Dietetic Association.* **79(5)** 542–547.

Centers for Disease Control (CDC), Food Safety Home. 2006. Outbreak Investigations, http://www.cdc.gov/foodsafety/outbreaks.htm (accessed Feb. 10, 2006).

__________________. 2003. Foodborne Diseases Active Surveillance Network (FoodNet) Emerging Infections Program Report on Foodborne Pathogens. http://www.cdc.gov/foodnet/pub/publications/2005/FNsurv2003.pdf (accessed Feb. 1, 2006).

Clemen, R. T. 1996. *Making Hard Decisions: An Introduction to Decision Analysis.* Brooks/Cole Pub. Co., Pacific Grove, CA.

Cooke, R. M. 1991. *Experts in Uncertainty: Opinion and Subjective Probability in Science* Oxford University Press, Oxford, UK.

Cooke, R., L. H. J. Goossens. 2004. Expert Judgement Elicitation for Risk Assessment of Critical Infrastructures. *J. of Risk Research* **7** 643–656.

________________________________. Undated. Expert Judgement Elicitations for Risk Assessments of Critical Infrastructures. Unpublished manuscript. Delft University of Technology, Delft, The Netherlands.

Craig, K. 1943. *The Nature of Explanation.* Cambridge University Press, Cambridge, UK.

DeWaal, C. S., K. Barlow. 2002. Outbreak Alert! Closing the Gaps in Our Federal Food-Safetynet. Center for Science in the Public Interest, Washington, D.C.

Epley, N., T. Gilovich. 2001. Putting Adjustment Back into the Anchoring-and-Adjustment Heuristic: Self-generated versus Experimenter Provided Anchors. *Psychological Science* **12** 391–396.

Gentner, D., A. Stevens (eds). 1983. *Mental Models*. Lawrence Erlbaum Associates: Hillsdale, New Jersey.

George, J. F., K. Duffy, M. Ahuja. 2000. Countering the Anchoring and Adjustment Bias with Decision Support Systems. *Decision Support Systems.* **29** 195–206.

Hoffmann, S. A., P. Fischbeck, M. McWilliams, A. Krupnick. 2006. Using Expert Elicitation to Link Food Consumption to Foodborne Illnesses. Unpublished manuscript. Resources for the Future, Washington, DC.

Keeney, R. L., D. von Winterfeldt. 1991. Eliciting Probabilities from Experts in Complex Technical Problems. *IEEE Transactions on Engineering Management* **38(3)** 191–201.

Keith, D. W. 1996. When Is It Appropriate to Combine Expert Judgements? *Climatic Change* **33** 139–143.

Lempert, R. J., N. Nakicenovic, D. Sareqitz, M. Schlesinger. 2004. Characterizing Climate Change Uncertainties for Decision-makers. *Climate Change* **45** 1–9.

Mead, P. S.,et al. 1999. Food-Related Illness and Death in the United States. *Emerging Infectious Disease* **5** 607–625.

Merrick, J. R. W., J. R. van Dorp, A. Singh. 2005. Analysis of Correlated Expert Judgments from Extended Pairwise Comparisons. *Decision Analysis* **2(1)** 17–29.

Morgan, M. G., B. Fischhoff, A. Bostrom, C. J. Atman. 2002. *Risk Communication: A Mental Models Approach.* Cambridge University Press, Cambridge, MA.

Morgan, M. G., L. F. Pitelka, and E. Shevliakova. 2001. Elicitation of Expert Judgments of Climate Change Impacts on Forest Ecosystems. *Climatic Change* **49** 279–307, 2001.

Morgan, M. G., et al. 2000. Categorizing Risks for Risk Ranking. *Risk Analysis* **20** 49–58.

Morgan, M. G., M. Henrion, 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis.* Cambridge University Press, Cambridge, MA.

Rea, L. M., R. A. Parker. 1997. *Designing and Conducting Survey Research: A Comprehensive Guide*. Jossey-Bass Publishers, San Francisco, CA.

Resources for the Future. March 13–14, 2006. Expert Judgment Policy Symposium and Technical Workshop. Organized in conjunction with Harvard Center for Risk Analysis, and Carnegie Mellon's Department of Engineering and Public Policy. Washington, D.C.

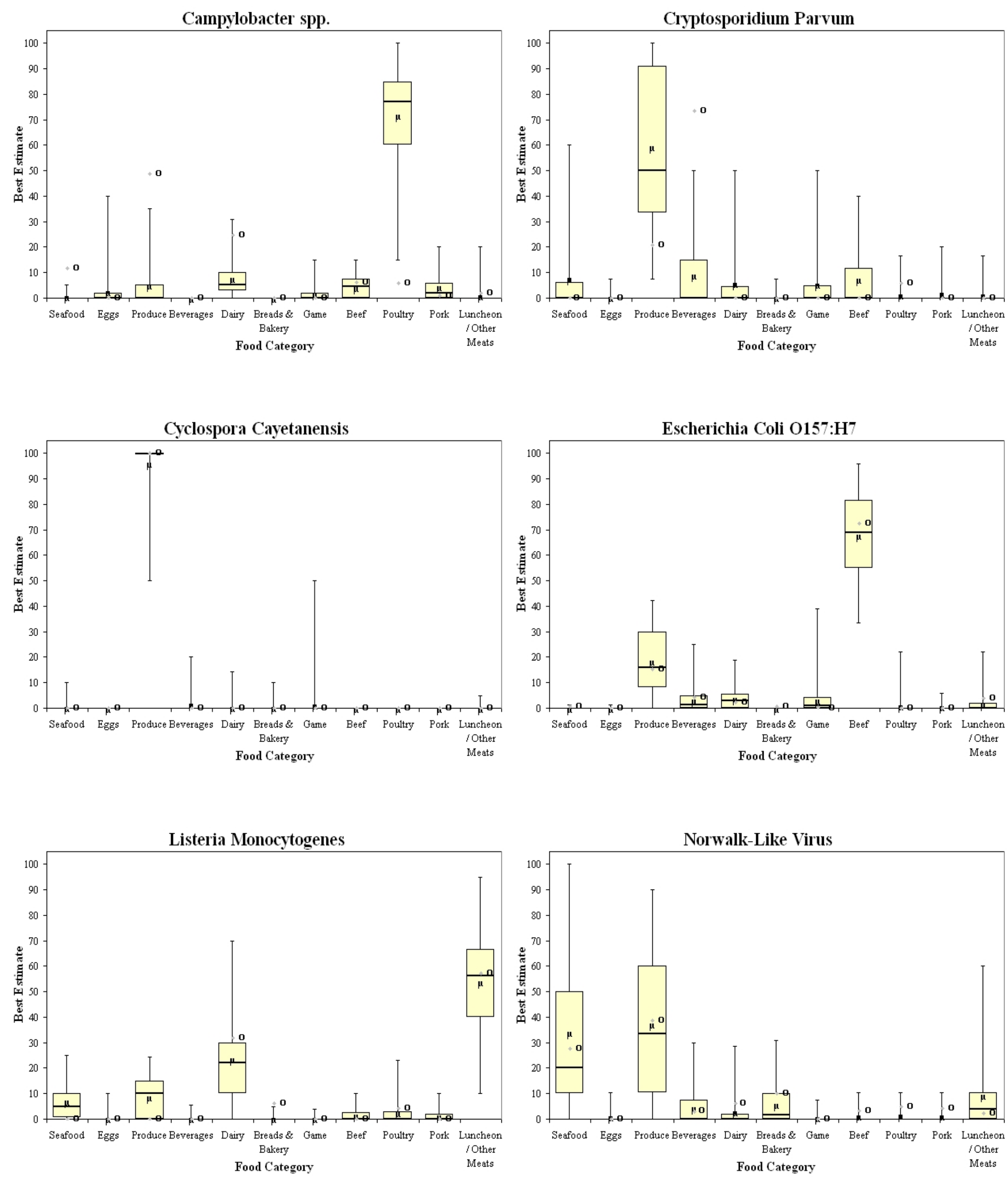
Tversky, A., D. Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* **185** 1124–1131.

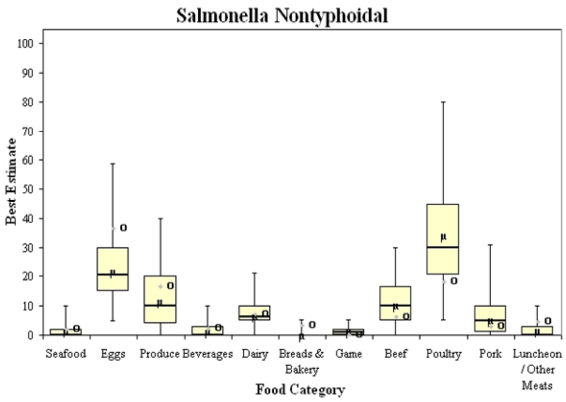
Van der Fels-Klerz, I. J. H., L. H. J. Goossens, H. W. Saatkamp, S. H. S. Horst. 2002. Elicitation of Quantitative Data from a Heterogeneous Expert Panel: Formal Process and Application in Animal Health. *Risk Analysis* **22(1)** 67–81.

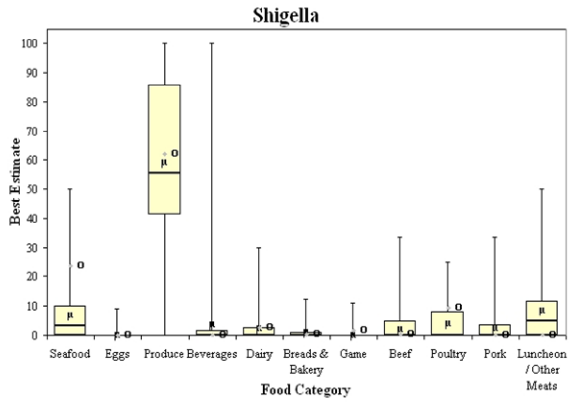
Winkler, R. L., R. T. Clemen. Multiple Experts vs. Multiple Methods: Combining Correlation Assessments. *Decision Analysis* **1(3)** 167–176.

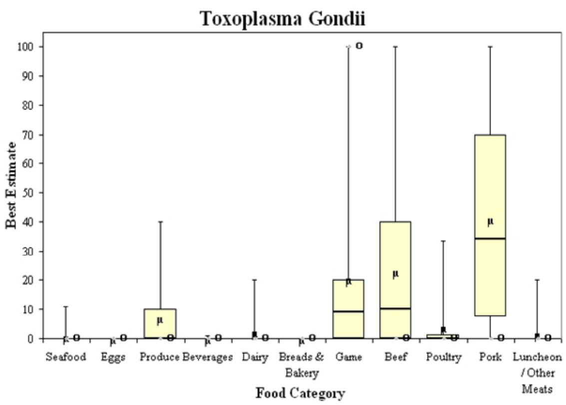
Figures and Tables

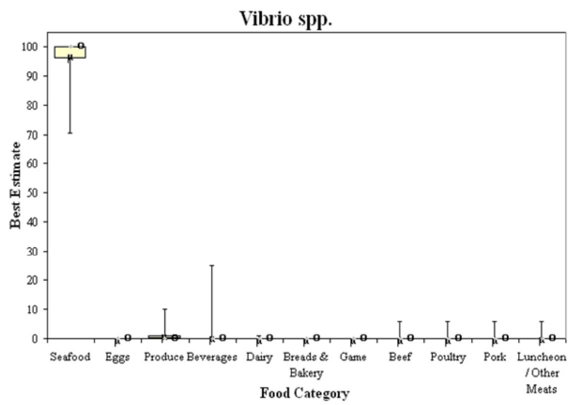
Figure 1. Box and Whisker Diagrams of Expert Elicitation Best Estimates of Foodborne Illness Food Attributions in the U.S. in a Typical Year

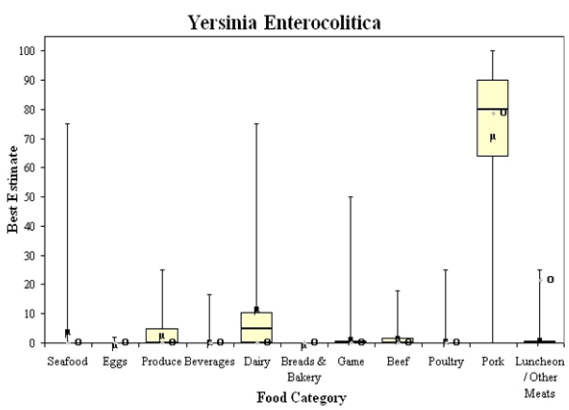












NOTE: μ denotes mean expert best attribution estimate. O denotes outbreak-based attribution estimate.

Figure 2. Instruction Sheet for Expert Respondents from Expert Elicitation Foodborne Illness Attribution Survey

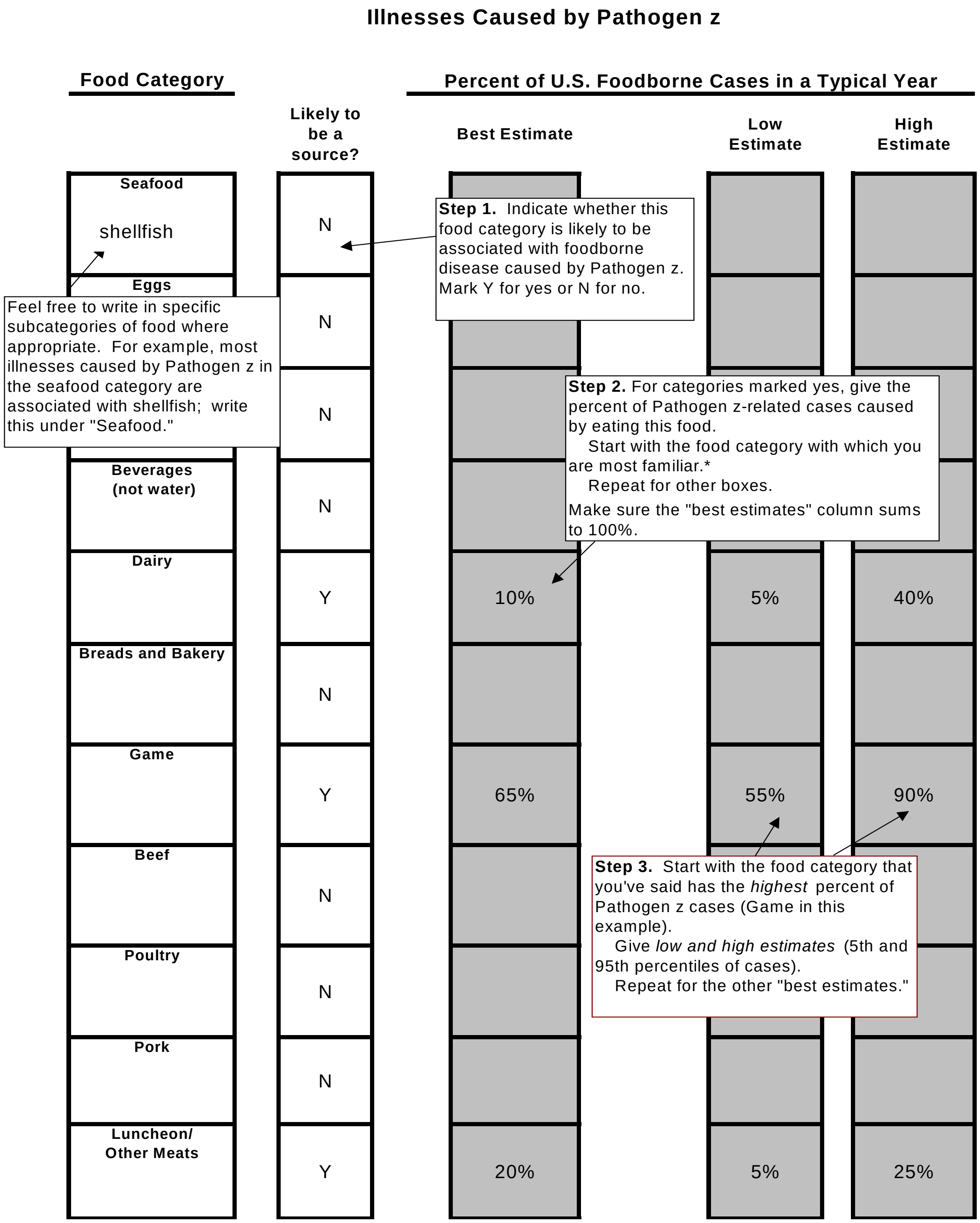
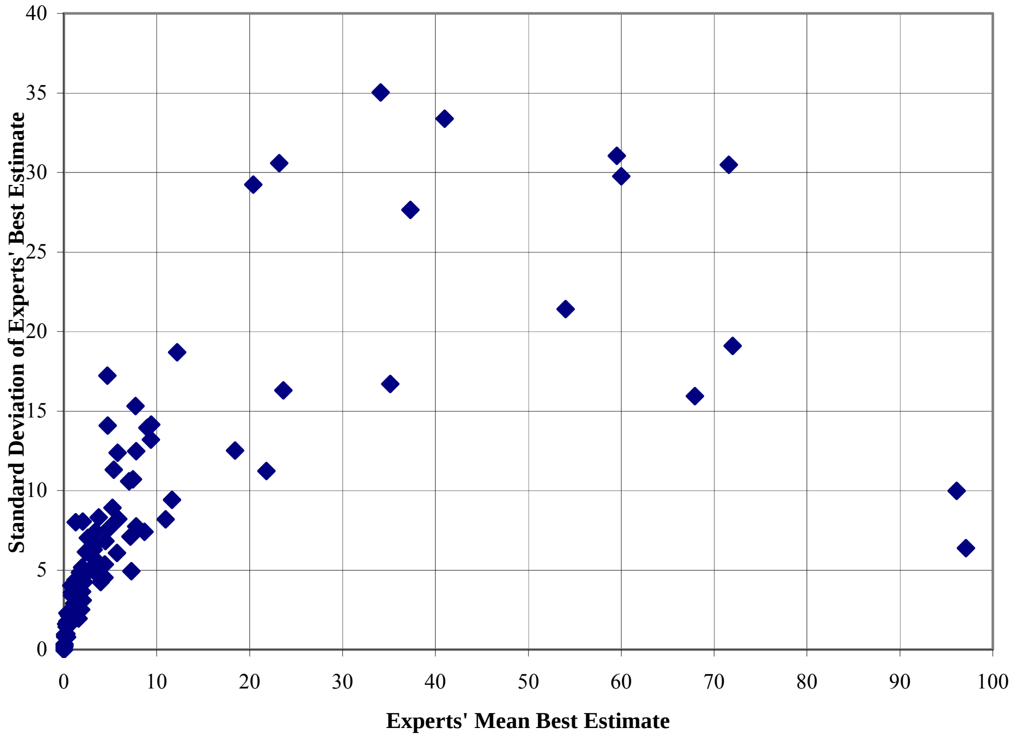
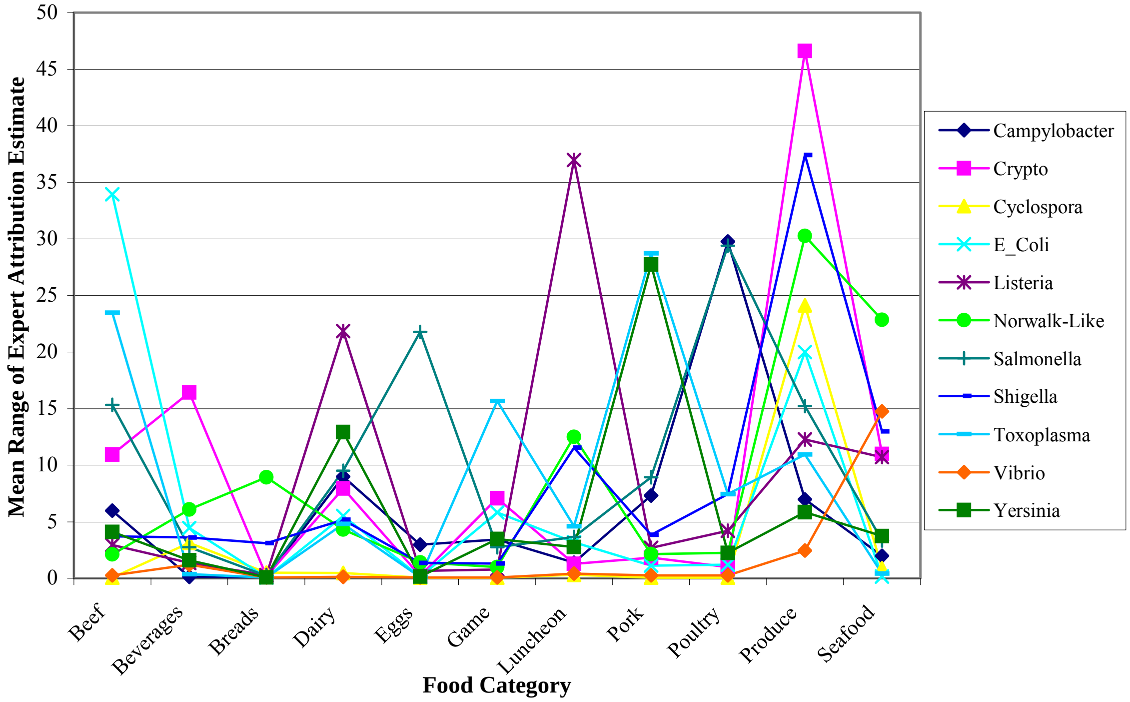


Figure 3. Variability in Experts' Best Estimate Food Attribution Judgments Relative to Mean Best Estimate



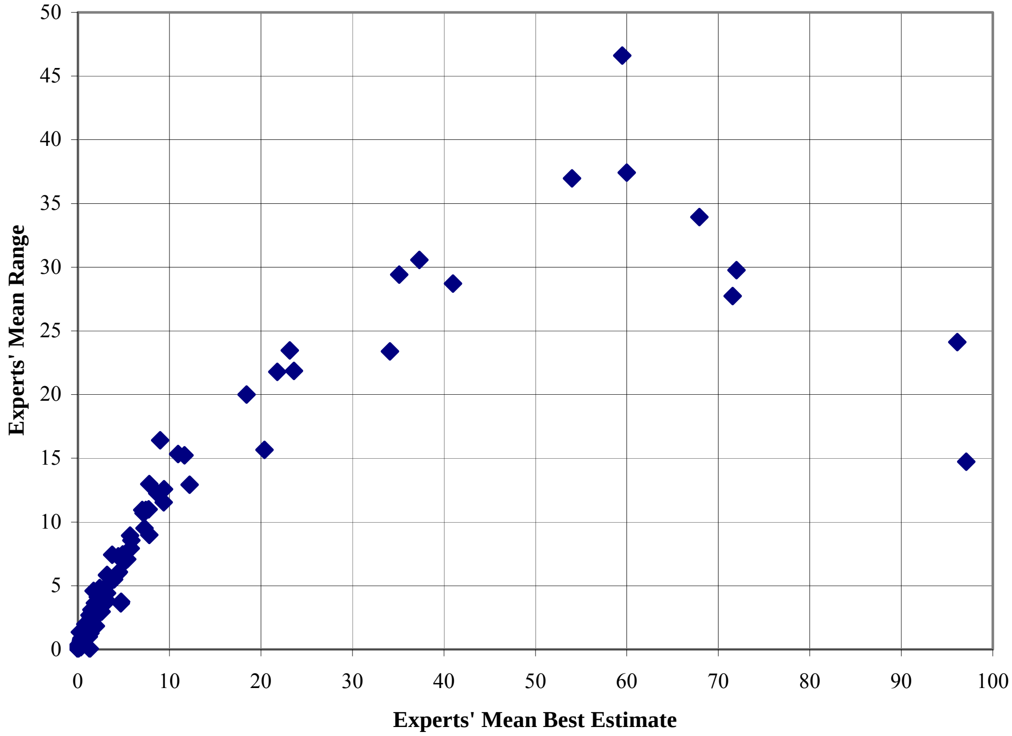
NOTE: Each point represents one of 121 distinct combinations of the 11 pathogens and 11 food categories.

Figure 4. Mean Range of Expert Attribution Estimates by Pathogen and Food Category



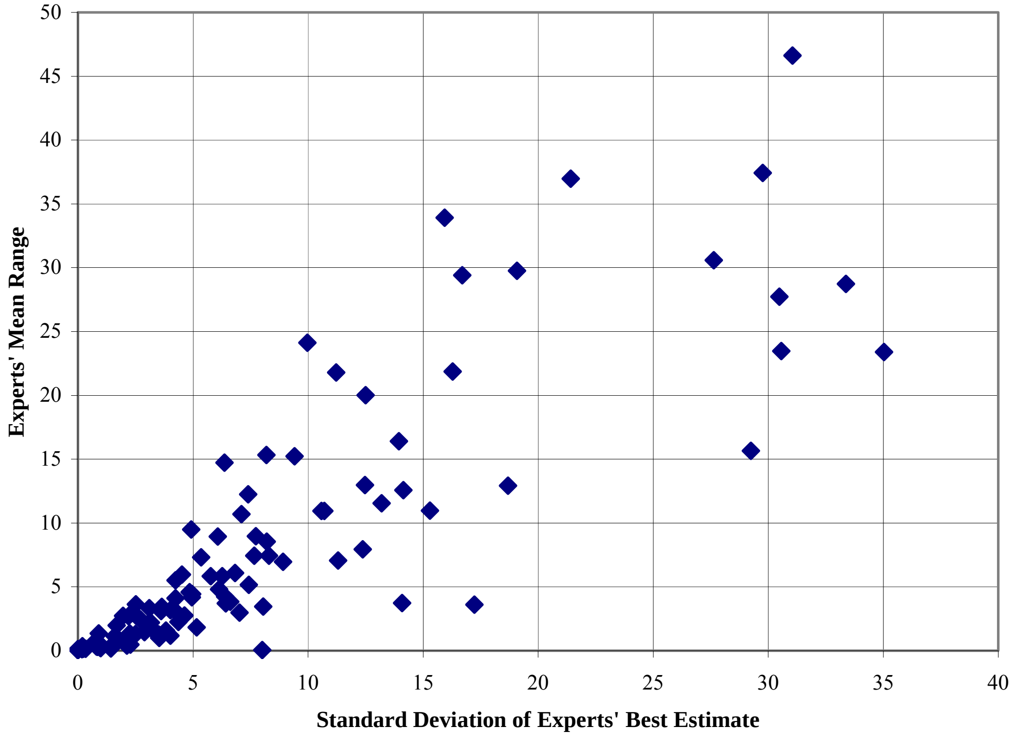
NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval.

Figure 5. Mean Range of Expert Foodborne-Illness Attribution Estimates Relative to Expert Mean Best Estimates



NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval Each point represents one of 121 distinct combinations of the 11 pathogens and 11 food categories.

Figure 6. Experts' Mean Range Relative to the Variability in Best Estimates Across Experts



NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval. Each point represents one of 121 distinct combinations of the 11 pathogens and 11 food categories.

Figure 7. A Framework Using Uncertainty Comparisons to Inform Risk Management Decisions

Individual uncertainty	Variabilty between Experts: low	Variabilty between Experts: high
low	Certainty & agreement	Certainty & disagreement
high	Uncertainty & agreement	Uncertainty & disagreement

Expert Agreement with Outbreak Attribution	Variabilty between Experts: low	Variabilty between Experts: high
close	Experts agree with outbreak attribution.	Unlikely
far	Experts don't think outbreak data is full story.	Experts don't agree on what drives attribution.

Table 1. Self-Evaluated Food and Pathogen Expertise (Scored with a 5-point Likert Scale: 1 low and 5 high)

Food Categories

	Mean	St. Dev.
Beef	3.69	0.95
Poultry	3.53	0.99
Eggs	3.2	0.97
Produce	3.18	1.17
Pork	3.16	1.15
Dairy	3.13	1.04
Seafood	3.13	1.16
Lunch Meat	3.09	1.22
Beverages	2.24	0.98
Breads & Bakery	2.22	1.04
Game	2	0.93

Pathogens

	Mean	St. Dev.
Escherichia Coli O157:H7	3.89	0.78
Salmonella	3.73	0.95
Listeria	3.65	1.03
Campylobacter	3.16	1.06
Vibrio	2.89	1.15
Shigella	2.75	1.01
Norwalk-like viruses	2.64	1.31
Yersinia	2.64	1.08
Cryptosporidium	2.24	1.07
Cyclospora	2.04	1.09
Toxoplasma	1.98	1.03

Table 2. Statistical Comparison of Expert and Outbreak-based Food Attributions

Pathogen	Food Category	Outbreak-Based Estimate	Mean Best Estimate	Standard Error	Mean Range*	% of Experts' Ranges Not Including Outbreak Est.	z-Test of Ho: Outbreak Est. = Mean Expert Est.
Campylobacter							
	Beef	6.1	4.4	0.73	6.0	48.4%	-2.39 **
	Dairy	24.6	7.8	1.26	9.0	93.9%	-13.38 ***
	Poultry	5.8	72.0	3.10	29.8	100.0%	21.36 ***
	Produce	48.9	5.2	1.45	7.0	93.9%	-30.20 ***
	Seafood	11.8	0.8	0.27	2.0	97.1%	-40.23 ***
Cryptosporidium							
	Beef	0.0	7.4	1.76	10.9	32.3%	4.22 ***
	Beverages	73.5	9.0	2.29	16.4	93.8%	-28.12 ***
	Dairy	0.0	5.8	2.04	7.9	32.3%	2.84 ***
	Game	0.0	5.4	1.86	7.1	20.0%	2.89 ***
	Poultry	5.8	1.2	0.58	1.0	96.8%	-7.88 ***
	Produce	20.8	59.5	5.11	46.6	61.3%	7.59 ***
	Seafood	0.0	7.7	2.52	11.0	28.1%	3.07 ***
Cyclospora							
	Produce	100.0	96.1	1.60	24.1	11.8%	-2.43 **
Escherichia Coli O157:H7							
	Beef	72.6	67.9	2.49	33.9	31.4%	-1.89 *
	Produce	15.4	18.4	1.95	20.0	50.0%	1.56
Listeria							
	Breads	6.3	0.2	0.14	0.3	97.1%	-44.91 ***
	Dairy	32.0	23.6	2.61	21.9	55.9%	-3.22 ***
	Lunch Meat	57.3	54.0	3.43	37.0	41.2%	-0.97
	Produce	0.0	8.7	1.19	12.3	60.6%	7.33 ***
	Seafood	0.0	7.2	1.14	10.7	61.8%	6.29 ***
Norwalk-like Viruses							
	Breads	9.9	5.8	1.43	8.5	66.7%	-2.87 ***
	Dairy	6.1	2.9	1.11	4.2	76.7%	-2.90 ***
	Lunch Meat	2.2	9.4	2.46	12.6	83.3%	2.92 ***
	Produce	38.8	37.3	4.81	30.6	70.0%	-0.31
	Seafood	27.7	34.1	6.10	23.4	70.0%	1.06
Salmonella							
	Beef	6.3	10.9	1.37	15.3	47.1%	3.41 ***
	Dairy	7.2	7.3	0.82	9.5	35.3%	0.07
	Eggs	36.5	21.8	1.87	21.8	70.6%	-7.85 ***
	Pork	2.9	5.7	1.01	8.9	58.8%	2.73 ***
	Poultry	18.3	35.1	2.78	29.4	70.6%	6.06 ***
	Produce	16.6	11.7	1.57	15.2	44.1%	-3.15 ***
Shigella							
	Lunch Meat	0.0	9.4	2.23	11.6	50.0%	4.20 ***
	Poultry	9.3	4.9	1.30	7.4	81.3%	-3.37 ***
	Produce	62.0	60.0	5.03	37.4	43.3%	-0.39
	Seafood	23.6	7.8	2.11	13.0	87.5%	-7.50 ***
Toxoplasma							
	Beef	0.0	23.2	5.40	23.5	59.3%	4.29 ***
	Game	100.0	20.4	5.17	15.7	96.3%	-15.40 ***
	Pork	0.0	41.0	5.90	28.7	74.1%	6.95 ***
	Produce	0.0	7.0	1.87	11.0	40.7%	3.76 ***
Vibrio							
	Seafood	100.0	97.1	1.02	14.7	3.0%	-2.84 ***
Yersinia							
	Dairy	0.0	12.2	2.99	12.9	63.6%	4.08 ***
	Lunch Meat	21.4	1.8	0.74	2.8	94.1%	-26.50 ***
	Pork	78.6	71.6	4.88	27.7	38.2%	-1.43

*** Statistically different from zero at the 1% level

** Statistically different from zero at the 5% level

* Statistically different from zero at the 10% level

NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval.

Table 3. Tobit Regression Results for the Difference between Mean Best Estimate of the Expert Elicitation and Outbreak-Based Food Attribution

Explanatory Variables	I	II	III	IV	V
Industry			-0.98	-0.25	-.28
Academia			0.08	0.03	-0.35
Multisector			0.42	0.07	-0.28
PhD			-.36	-.06	0.1
DVM					
MS			-.11	.00	0.05
Veterinary Medicine			-.88	omitted	omitted
Public Health			-.02	.03	0.25
Human Medicine			-.95	-.11	0.31
Food Science			-.98	-.07	-0.14
Other Field			-.28	0.07	0.11
Pathogen Expertise			-.58	-0.03	-0.01
Food Expertise			1.35***	0.11	0.13
Years of Experience			-.034	-0.01	-.01
No. of Outbreaks			-.047***	-.08***	-.08***
Campylobacter	omitted			omitted	omitted
Norwalk-like viruses	-3.66***			0.47	0.49
Salmonella.	-3.7***			3.49	3.48
Cryptosporidium	-6.03***			-0.45	-0.43
Shigella	-6.19***			-0.11	-0.09
Toxoplasma	-6.43***			-0.35	-0.33
E. Coli O157:H7	-10.37***'			0.62	0.61
Listeria	-11.23***			-0.17	-0.18
Yersinia	-16.15***			-0.24	-0.23
Vibrio	-41.39***			-0.01	0
Cyclospora	-45.01***			-0.22	-0.19
Produce		omitted		omitted	omitted
Poultry		-8.67***		6.4***	6.49***
Seafood		-11.28***		-0.86	-0.85
Dairy		-11.57***		-1.33	-1.33
Lunch Meat		-12.16***		-2.07***	-2.06*
Beef		-13.18***		2.39**	2.38**
Pork		-14.05***		2.44**	2.44**
Game		-14.67***		-5.95***	-5.91***
Beverage		-15.5***		-6.18***	-6.15***
Breads		-22.35***		-1.96	-1.93
Eggs		-34.75***		-0.36	-0.36
Expert ID Dummies					
No. obs.	4483	4483	4483	4483	4483
LR chi2	926.12	381.54	54.46	220.18	220.38
pr > chi2	0	0	0	0	0

*** Statistically different from zero at the 1% level
** Statistically different from zero at the 5% level
* Statistically different from zero at the 10% level

Table 4. Top Decile Rankings of Expert Elicitation Food-Pathogen Combinations with the Greatest Variability in Foodborne Illness Attributions

Pathogen	Food Category	Absolute Value of Difference Between Expert Best Est. & Outbreak Est. (percentage points)		Average Range (percentage points)		Standard Deviation of Experts' Best Estimates (percentage points)	
Campylobacter	Poultry	66.14	(2)	29.76	(6)	19.09	(10)
Toxoplasma	Pork	41.00	(5)	28.72	(8)	33.37	(2)
Cryptosporidia	Produce	38.74	(6)	46.62	(1)	31.06	(3)
Toxoplasma	Beef	23.17	(7)	23.47	(11)	30.57	(4)
Salmonella	Poultry	16.87	(9)	29.41	(7)	16.70	(13)
Shigella	Produce			37.42	(2)	29.76	(6)
Listeria	Lunch Meat			36.97	(3)	21.41	(9)
Norwalk	Produce			30.58	(5)	27.64	(8)
Yersinia	Pork			27.73	(9)	30.49	(5)
Norwalk	Seafood			23.39	(12)	35.03	(1)
Toxoplasma	Game	79.61	(1)			29.24	(7)
Yersinia	Dairy	12.20	(13)			18.69	(11)
E.Coli O157:H7	Beef			33.93	(4)		
Cyclospora	Produce			24.12	(10)		
Listeria	Dairy			21.86	(13)		
Cryptosporidia	Beverages	64.48	(3)				
Campylobacter	Produce	43.67	(4)				
Yersinia	Lunch Meat	19.66	(8)				
Campylobacter	Dairy	16.80	(10)				
Shigella	Seafood	15.82	(11)				
Salmonella	Eggs	14.69	(12)				
Shigella	Beverages					17.23	(12)

NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval.

Table 5. Tobit Regression Results for the Standard Deviation of the Expert Elicitation Best Estimate Food Attribution

Explanatory Variables	I	II	III
Mean Best Estimate	1.22***	1.14***	1.14***
Mean Best Estimate Squared	-.01***	-.01***	-.01***
No. of Outbreaks	-.05***	-.04***	-.04***
Pathogen Expertise	-1.99***		-2.13***
Food Expertise	-.68		0.87
Listeria		-0.46	-0.63
E. coli O157:H7		-0.01	0.33
Campylobacter		0.58	-0.63
Vibrio		0.61	-1.18
Cyclospora		1.29	-2.3**
Norwalk-like Viruses		3.14***	0.82
Cryptosporidium		3.65***	0.5
Toxoplasma		3.73***	dropped
Yersinia		3.73***	1.41
Shigella		3.76***	1.68*
Salmonella.		omitted	omitted
Eggs		0.8	-0.04
Poultry		0.86	-0.27
Beef		1.27	dropped
Lunch Meat		1.66	0.91
Produce		1.73	0.9
Dairy		2.01*	1.22
Pork		2.24*	1.43
Beverages		2.88**	2.86
Game		3.41***	3.6***
Seafood		3.94***	3.15***
Breads		omitted	omitted
No. obs.	121	121	121
LR chi2	241.36	277.02	277.02
pr > chi2	0	0	0

*** Statistically different from zero at the 1% level
** Statistically different from zero at the 5% level
* Statistically different from zero at the 10% level

Table 6. Expert Elicitation Mean Range of Best Estimate Foodborne Illness Attribution by Pathogen and Food Category

Pathogen	Mean Range	Food Category	Mean Range
Norwalk-like viruses	10.9	Produce	19.0
Shigella	9.0	Beef	9.3
Cryptosporidium	8.7	Poultry	7.8
Listeria	8.2	Dairy	7.5
Salmonella	7.5	Pork	7.4
Vibrio	6.5	Seafood	7.4
Yersinia	6.5	Lunch Meat	7.2
Campylobacter	6.3	Beverages	3.7
Cyclospora	5.7	Game	3.6
Escherichia Coli O157:H7	3.8	Eggs	2.7
Toxoplasma	2.2	Breads	1.1

NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval.

Table 7. Tobit Regression Results for the Mean Range of Expert Elicitation Food Attribution Estimates

Explanatory Variables	I	II	III	IV
Best Estimate	1.45***	1.47***	1.46***	1.42***
Best Estimate Squared	-.01***	-.01***	-.01***	-.01***
Government	omitted		omitted	omitted
Academia	2.32***		2.43***	2.35***
Industry	4.64***		4.37***	4.72***
Multisector	11.05***		12.34***	12.61***
DVM		-6.8***	omitted	omitted
MD		-3.35***	omitted	omitted
PhD		-3.25***	omitted	omitted
MS		omitted	omitted	omitted
Microbiology			omitted	omitted
Public Health			4.63***	4.08***
Food Science			2.29***	2.29***
Human Medicine			-.08	-.64
Other Field			-.26	-.70
Veterinary Medicine			-2.95***	-3.53***
Pathogen Expertise			-1.27***	-.76***
Food Expertise			-.03	-.15
Campylobacter				omitted
Cryptosporidium				5.82***
Toxoplasma				4.12***
Vibrio				-2.41
Shigella				1.40
E. Coli O157:H7				-1.19
Yersinia				-1.14
Norwalk-like viruses				0.79
Cyclospora				0.92
Salmonella				0.60
Listeria				-0.49
Produce				omitted
Dairy				-4.1***
Poultry				-4.96***
Beef				-3.01***
Pork				-2.86***
Game				-2.58**
Seafood				-2.15**
Breads				-3.07*
Eggs				-3.06*
Lunch Meat				-1.78*
Beverages				-1.43
Years of Experience			-..09***	-.08**
No. of Outbreaks			-.02***	-.02*
No. obs.	1566	1566	1566	1566
LR chi2	1420.18	1383.12	1526.44	1954.74
pr > chi2	0	0	0	0

*** Statistically different from zero at the 1% level
** Statistically different from zero at the 5% level
* Statistically different from zero at the 10% level

NOTE: Experts were requested to provide 90 percent confidence intervals around each best estimate of foodborne illness food attribution. Range is equal to the size of the confidence interval.