

Attributing Foodborne Illness Using Consumption Data and Expert Elicitation

Amber Jessup (U.S. Department of Health and Human Services)*, Aylin Sertkaya (Eastern Research Group, Inc), Kara Morgan (U.S. Food and Drug Administration)

Working Paper prepared for:

Methods for Research Synthesis: A Cross-Disciplinary Workshop

Harvard Center for Risk Analysis

October 3, 2013

www.hcra.harvard.edu

***Corresponding author:** amber.jessup@hhs.gov

Disclaimer: The findings and conclusions of this paper are those of the authors and do not imply endorsement by any component of Harvard University or other sponsors of this workshop, nor do they necessarily represent the views of the U.S. Department of Health and Human Services. Comments should be directed to the authors.

Acknowledgment: The funding for this study was provided by the U.S. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation (ASPE).

ATTRIBUTING FOODBORNE ILLNESS USING CONSUMPTION DATA AND EXPERT ELICITATION

Amber Jessup (U.S. Department of Health and Human Services)*, Aylin Sertkaya (Eastern Research Group, Inc), Kara Morgan (U.S. Food and Drug Administration)

ABSTRACT

For regulators, understanding what foods are likely to be contaminated with pathogens, the likelihood that these foods will cause illness, and where in the supply chain that the contamination is likely to be introduced are critical to implementing effective risk control and making efficient use of scarce food safety resources. To understand the likelihood of illness associated with microbiological contamination of foods close to the manufacturing stage, before consumer or food service handling, this study focuses on the likelihood of illness associated with foods as they sit on the store shelf. To address the challenges of lack of reliable data at sufficient degree of granularity and difficulties in eliciting expert opinion on attribution rates for a large number of food products, we develop and apply an innovative expert elicitation protocol to generate risk rankings for highly disaggregated food categories. Using these risk rankings, we will be able to combine the output of the expert elicitation with data on consumption shares to estimate the number of foodborne illness cases by food category. The method we employ better incorporates the knowledge and experience of the experts, and reduces the judgment biases. In addition to the data generated from the expert elicitation, this study relies on sales data on foods consumed in the United States available from The Nielsen Company and select food trade associations, such as the International Dairy Deli Bakery Association (IDDBA). This study addresses the challenges in understanding the relative contribution of different types of food to foodborne illness. Expert elicitation provides a method to provide estimates, and for this situation, expert elicitation provides an approach for estimating model parameters and factors that are of interest. In cases, such as those for this study, expert elicitation may be the best, or only, way to obtain such estimates.

Keywords: expert elicitation, foodborne illness, attribution, food consumption data, microbiological risk

1 INTRODUCTION

Regulatory tools are sometimes most efficiently applied at intermediate steps, before food reaches the consumer. For example, food safety regulatory agencies can intervene at the point of manufacturing, packaging or labeling, when food is entering the United States, or while it is being held at a warehouse or in a retail outlet. For this study, we were interested in estimating the likelihood of contamination close to the manufacturing stage. In other attribution studies, because of the epidemiological data that these studies relied on, the likelihood of causing illness was assessed at the point of consumption. The most effective targeting of intervention can only be identified if the point in the supply chain at which the foods are contaminated is known. To understand the likelihood of illness associated with contamination of foods close to the manufacturing stage, that is, before consumer or food service handling, this study focuses on the likelihood of illness associated with foods as they sit on the store shelf.

To address the challenges of the lack of reliable data and difficulties in applying expert elicitation, we develop and apply an innovative expert elicitation protocol to generate risk rankings for highly disaggregated food categories. Using these risk rankings, we will be able to combine the output of the expert elicitation with data on consumption shares to estimate the number of foodborne illness cases by food category. Our premise is that the method employed better incorporates the knowledge and experience of the experts, and reduces the judgment biases.

Policy analysts are often asked to answer questions for which there are (1) no peer-reviewed studies, (2) several apparently conflicting peer-reviewed studies, or (3) peer-reviewed studies that are only peripherally related to the policy question. Meta-analyses, systematic reviews and expert elicitation are all possible approaches to synthesizing information under these circumstances. Meta-analysis and systematic reviews synthesize the existing literature and, thus require a relatively sizable body of published studies. Expert elicitation, on the other hand, synthesizes expert opinions when data of sufficient quality are lacking, and can also be used to characterize uncertainty. Therefore, for situations in which existing data and studies are insufficient, expert elicitation is the most viable of these three options.

Analyzing the relative likelihood of foods causing foodborne illness is an area well suited to expert elicitation. Although there are well-established surveillance systems and analytical methods that provide estimates of the **total** number of illness in the United States each year from pathogens that have been associated with food (Scallan, et al., 2011; Scallan, Griffin, Angulo, Tauxe, & Hoekstra, 2011), estimates

linking these illnesses to specific foods, as opposed to aggregate food groups (e.g., produce, dairy, meat and poultry, etc.) are lacking. Existing studies estimate number of foodborne illnesses for food groupings that are highly aggregated and heterogeneous with respect to microbiological contamination risks. These studies are also largely dependent on outbreak data, which only represent about 5 percent of foodborne illnesses and may not be representative of all foodborne illnesses.

However, expert judgments are known to be influenced by heuristics (Tversky & Kahneman, 1974) that can lead to biases in results, and even the most structured expert elicitation process is not always sufficient to overcome those biases. For example, expert opinions on previous expert elicitation studies could be affected by the availability bias and the anchoring bias if opinions are heavily influenced with recent known foodborne illness outbreaks.

Deriving estimates of the likelihood of specific foods causing foodborne illness is an important area of study because the burden from foodborne illness in the United States is significant. Each year 14 major foodborne pathogens are estimated to cause approximately 9 million illnesses, resulting in the loss of over 61,000 QALYs and 14.1 billion dollars in cost of illness (Batz, Hoffman, & Morris, 2011). In spite of the public health significance of foodborne illness, information needed for regulatory agencies to target food safety policy is lacking.

Understanding what foods are likely to be contaminated, the likelihood these foods will cause illness, and where in the supply chain that the contamination was introduced is critical to implementing effective risk control and making efficient use of scarce food safety regulatory resources. However, linking illnesses to specific foods is very challenging. Foodborne illnesses may occur within hours or as long as weeks after eating the contaminated food. Linking foodborne illness to the causal food vehicle typically only occurs in outbreaks, when epidemiologists use case-control studies to establish common foods consumed by affected consumers and trace back to the food vehicle. For almost all sporadic cases and even for almost half of known outbreaks, the food vehicle is never determined.

Painter et al. (2013) use outbreak data to attribute foodborne illnesses to 17 food commodities. They find that 46 percent of illnesses are linked to produce and the most deaths are linked to poultry. Although data from foodborne disease outbreaks provides a readily available source of information, the reliability and availability of this data in estimating the number of foodborne illnesses in the population attributable to each food source is limited. Outbreak-associated illnesses represent a small subset of all foodborne

illnesses, and illnesses reported in outbreak data may not be representative of all illnesses because of factors related to the characteristics of food contamination, food preparation and service, pathogen transmission, and the ease of outbreak detection and investigation for certain foods and pathogens. Outbreak data are also limited by missing and incomplete data fields, resulting in a high level of uncertainty and potential errors in foodborne illness attribution estimates. Likewise, data may be biased toward large outbreaks, outbreaks associated with point sources, outbreaks that have short incubation periods, and outbreaks that cause serious illness. For example, the Painter et al (2013) study was not able to estimate attribution for *Toxoplasma* spp. or *Vibrio Vulnificus*, because no outbreaks were identified linked to them, even though these are known foodborne pathogens. These limitations make estimates of the number of illnesses associated with food-pathogen pairs based on only outbreak data highly uncertain.

Other foodborne illness attribution studies (Batz, Hoffman, & Morris, 2011; Painter, et al., 2013; Pires, et al., 2009; Karns, Muth, & Cogliaiti, 2005) use expert elicitation to estimate the number of illnesses by food categories. They are limited, however, by estimating attribution for only for a small number of highly aggregated food categories.

These studies, although helpful in identifying areas of potential concern, do not provide sufficiently detailed groupings of foods to meet the food safety regulatory agency needs. Placing all foods into 10 to 20 categories does not allow for sufficient homogeneity with regard to microbiological contamination risk within the categories. For example, frozen or canned produce would have a very different risk profile than raw produce. Putting all of these heterogeneous categories of risk into one overarching category (i.e., produce) limits regulatory agencies' ability to identify the foods of highest risk, and complicates risk ranking activities. Attribution of foodborne illness to groupings of foods that are homogeneous with respect to microbiological contamination risk would lead to a better understanding of the best opportunities for regulatory agencies to target their food safety resources.

2 DATA

In addition to the data generated from the expert elicitation, this study relies on sales data on foods consumed in the United States available from The Nielsen Company and select food trade associations, such as the International Dairy Deli Bakery Association (IDDBA). The Nielsen Company data included information on total dollar sales, total sales units, and total equivalent unit sales (which are total sales

reported on an equivalized basis, such as pounds, cases, servings, etc.) for U.S. food, drug, and mass merchandisers for 52 weeks ending September 5, 2009 at the Universal Product Code (UPC) level. Detailed Nielsen scanner data was not available for eight of the food categories that were a part of the elicitation, namely: fresh meat, fresh poultry, fresh seafood, deli meat and deli cheese, other deli prepared foods, random-weight fresh produce, and prepared-in-store bakery items. Thus, supplementary total sales data for these categories was obtained from the IDDBA International Dairy Deli Bakery Association (IDDBA) 2011 annual trends report (“What’s in Store”). The IDDBA report provided comparable Nielsen scanner data total dollar sales for 52 weeks ending March 28, 2010.

As described in Section 3.3, the Nielsen Company data along with data on random weight products published by select trade associations were used to compute relative consumption shares. For the calculations, serving size for each food commodity group was needed. Assumptions about average serving sizes were combined with reported total equivalent unit sales information to estimate the relative total consumption in terms of total servings.

3 STUDY METHODOLOGY

This study uses an innovative approach for eliciting expert input to estimate foodborne illness attribution. Rather than a direct elicitation of attribution fractions, an input to a calculation was estimated by the experts, and will be combined with data on food consumption to develop attribution estimates. Expert elicitation was also used to identify subcategories of food that are homogenous in terms of microbiological contamination risk, and to estimate the proportion of illnesses from each pathogen that were due to contamination that occurs before the product reaches the end consumer, hereinafter referred to as source contamination.

3.1 Expert Selection

In the identification of experts, we sought a balance in the experts’ breadth of experience (by food industry) and professional affiliation. Given our specific interest in the likelihood of source contamination, we chose to emphasize expertise in microbiology and food safety over epidemiology. We recruited 16 experts to provide sufficient expertise across the range desired disciplines. Experts were ultimately selected based on both their professional experience related to the public health risks posed by different foods, as well as their willingness to participate in the process.

The final set of experts included participants from academia, government and industry. About 30 percent of the panel members were from academia and 25 percent were federal and state employees. Nearly two-thirds of the panel had microbiology as their primary field of expertise. Characteristics of the panel are shown in Table 1.

The experts worked independently over the course of 5 months, and were only aware of each other's work when they saw the disaggregated categories to be ranked. This will be explained in further detail later in this section.

3.2 Designing the Expert Elicitation

Addressing Known Judgment Heuristics

As mentioned earlier, expert judgments are known to be affected by heuristics or mental shortcuts. The two human judgment heuristics that were of most concern in this research were the availability bias and the anchoring bias. The availability bias is a mental shortcut people use to make probability judgments based on how readily examples come to mind rather than the true frequency. This would have an impact in this case by experts overweighting foods that have been recently associated with foodborne illness. The anchoring bias is the tendency for human judgment to overweight available information even when it is known to be uncertain. Existing studies on attribution that have used outbreak data have led to shared knowledge among relevant experts about which foods are more “risky,” even though outbreak data is known to be very uncertain. It was expected that if experts were asked for attribution fractions, they would anchor on those estimates and were likely to make insufficient adjustments to account for the known uncertainty.

Focusing the Elicitation on Fewer and More Relevant Categories

Because of the long list of pathogens of interest in Scallan et al. (2011), and the potentially long list of food commodities, a priority early on in the elicitation was to have the experts help to decrease the scope of the task. This was accomplished in multiple steps. In the first round of the elicitation, we asked the experts to indicate pathogens of most concern for a total of 96 subcategories of foods (e.g., cheese, milk, yogurt, etc.) which were organized into 24 groups conceptualized as how they might be found grouped in

conventional supermarket aisles (e.g., dairy, frozen foods, etc.). More specifically, for each of the 96 food subcategories, experts assessed the relative likelihood that the given food product type – while still on the supermarket shelf – could be contaminated by different types of pathogens (22 pathogens in total) at a level that could cause foodborne illness even if the product was properly handled by the consumer. Those food-pathogen pairs deemed irrelevant in terms of source contamination risk were eliminated from subsequent rounds of the elicitation based on the assumption that the fraction of illnesses that would be attributed to these food-pathogen pairs would have a negligible impact in the final calculations. A further culling of the food categories was accomplished in the second round of the elicitation, which asked experts to place each food-pathogen pair into bins separated by risk of microbiological contamination. Those food-pathogen pairs that experts rated as having a negligible risk of microbiological contamination were subsequently eliminated from consideration. As discussed in Section 3.3 below, the list of food-pathogen pairs deemed relevant for estimating number of foodborne illnesses associated with each pathogen was finalized after reviewing the risk rankings of experts in the second round and eliminating those pairs that had relatively low microbiological contamination risk.

In addition to eliminating very low-risk food categories, allowing the experts to guide the construction of the food categories also allowed us to capture the experts' input into what the right level of aggregation should be for this elicitation. If the categories cover a very heterogeneous set of items in terms of likelihood of contamination (like “dairy”), then ranking that category relative to others might vary depending on which item the expert is focusing on during the ranking. For example, if experts think of hard cheese, they might consider that category of food to have a relatively low likelihood of *Listeria* contamination, for example, but if they were thinking of “dairy” in terms of “soft unpasteurized cheese” would lead to a different ranking. Rather than allow for that variation, the expert elicitation was designed to re-define the scope of the ranking to create subcategories of food that were homogeneous with respect to likelihood of contamination. These experts, with training as food technologists and microbiologists, would know that soft unpasteurized cheese would have a higher likelihood of supporting the growth of bacteria than some other types of dairy products.

Therefore, initially, the experts were asked to identify a subset of the food category that had a different (either higher or lower) likelihood of contamination than the others in the category. So, when presented with the category of “produce,” for example, the experts had the option of specifying that canned produce had a different likelihood profile than fresh produce. All of these sub-categories were then ranked separately by all the experts in a later round.

With these concurrent tasks for reducing the list and expanding the list, the final list of food/pathogen pairs to be ranked remained rather large. To ease the cognitive burden, the elicitation tool used the conceptual model of “walking around a grocery store” to help the experts with the task. The elicitation tool allowed for food product categories to vary, sometimes significantly, by pathogen, as not all foods are equally likely to be contaminated by all pathogens of interest. We also incorporated skip logics and response-dependent question prompts to the tool to ease the burden on the experts (see discussion below for details on the types of questions posed).

Defining the Quantity Being Elicited

Although it may be difficult for an expert to assign an exact percentage of the total illnesses to the category of “soft pasteurized cheese,” it would not be conceptually difficult for an expert to express the **relative** likelihood of contamination of the soft pasteurized cheese as compared to other dairy products. As mentioned earlier, these were experts in food technology, microbiology and epidemiology, so they would have a basic understanding of the characteristics of the food that would be linked to likelihood of contamination and likelihood of growth. Again, asking for an exact number might be problematic, but the relative ranking of the categories would seem to be a rather straight-forward judgment for the experts to make, and would likely be somewhat consistent across experts, leading to a reliable result.

In the second round of the elicitation, the experts were asked to indicate the relative likelihood of contamination as the product is sitting on a store shelf on a scale of “negligible,” “low,” “medium,” or “high.” For those that were assigned a “medium” or “high” ranking, the experts were asked to divide those again into low/medium/high, leading to an 8-point scale. Therefore, the first round of the expert elicitation created (1) subcategories that were internally homogeneous with respect to risk and (2) a relative ranking of all the categories in terms of likelihood of contamination as they are sitting on the store shelf. This approach provided the greater level of detail that was needed and also helped to mediate the anchoring bias, by making the influence of other attribution studies based primarily on outbreak data less likely to bias the results.

As mentioned, expert elicitation has been used in foodborne attribution studies in the past. Expert elicitation was used in the Hoffmann et al. (2007) study to directly elicit the experts’ judgments on the fraction of illnesses caused by each type of food. Because the results for the project being described in

this paper was meant to inform regulatory activities closer to the manufacturing level, there was a need to separately attribute illnesses that were caused later in the supply chain, for example, due to mis-handling in storage or at retail. Batz et al. (2011) did provide an estimate for the source of the contamination, but they did so using a very different approach. Their “handling” estimate came from the illnesses that were attributed to “multi-ingredient foods,” which included foods like sandwiches and salads. The assumption in their paper was that these types of foods are very often assembled later in the supply chain, and therefore the illnesses attributed to them could be characterized as being due to “handling” rather than to source contamination. As explained later, we took a different approach but ended up with somewhat similar results.

To answer the need of this project, the study was designed to elicit the information on relative likelihood of different foods being contaminated with microbial pathogens as they sit on the store shelf, and separately elicit the proportion of illnesses from each pathogen that can be attributed to something that happens after the product leaves the store shelf. This question was answered by asking the experts what portion of the illnesses by pathogen were due to source contamination (that is, contamination of the product before it gets to the store shelf) as opposed to other types of contamination. Once this assessment was made, then, the foodborne illness attribution rates were to be estimated by combining these expert judgments about relative likelihood of contamination on the store shelf with data on consumption patterns, and setting aside those illnesses that were attributed to something besides source contamination.

Reducing the Cognitive Burden

Given the large number of food-pathogen pairs that were being assessed, we were very concerned about the cognitive burden being placed on the experts. We attempted to reduce the cognitive burden in a number of ways. First, we did extensive pre-testing of the expert elicitation and made changes to the protocol based on feedback from that pre-test. Second, we used the experts’ own judgments to reduce the size of the problem by eliminating food-pathogen pairs with negligible risk. Third, we used an innovative method for structuring the questions around the layout of a grocery store to ease the cognitive burden of thinking of all these different types of foods. Fourth, we accepted tradeoffs to simplify the questions, for example, using a simple scale (negligible/low/medium/high) to assess the relative likelihood of contamination rather than asking for specific numbers or relationships (e.g., A is 2x more likely to be contaminated than B.). Finally, we used multiple rounds in the elicitation over several months, so the

experts would not have to complete all the assessments into one short time period. We discuss the possible implications of these tradeoffs on the results in Section 5.3.

Aggregating the Expert Judgments

In addition to ranking each subcategory of food on a relative scale to indicate likelihood of contamination as it sits on a shelf, the experts were asked to rate themselves on their level of expertise in each food-pathogen combination. This self-rating was used to weight the ranking that was provided; a higher self-assessed expertise leads to a higher weight. This approach was included because we knew that some experts were more familiar with some pathogens and/or food products than others, so equal weighting seemed inappropriate.

We used a weighted average in aggregating the relative likelihood of contamination across all experts. More specifically, we computed a weighted average relative likelihood score, \bar{x}_w , (see Equation 1) for each food-pathogen pair where the weights are each individual's self-assessed confidence level as follows:

Equation (1)

$$\bar{x}_w = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i}$$

where w_i is the expert's self-assessed confidence level (0 through 4), x_i is the relative likelihood score for expert i , and N is the number of experts that provided a relative likelihood of contamination score for the food-pathogen pair in question.

3.3 Calculating Foodborne Illness Attribution from Relative Contamination Data

The relative rankings generated with the expert elicitation indicate which food category-pathogen combinations are more or less likely to result in foodborne illness due to contamination before the product reaches the store shelf. An additional step will be needed to generate the estimated number of foodborne illnesses associated with those combinations. Illnesses result from the contamination of a food with a given pathogen and the consumption of that food; for example, even those foods with a very high likelihood of contamination may account for very few foodborne illnesses if they are rarely eaten. Thus, to attribute foodborne illnesses to food-pathogen combinations, we will combine the contamination-based

relative rankings from experts with data on the relative share of consumption of these food categories. Our foodborne illness attribution methodology consists of five distinct steps, each of which is discussed in further detail below and illustrated through an example in Table 2 and Table 3.

Step 1 – Map Expert-defined Food Categories to Nielsen Scanner Data. With the elicitation, we identified a total of 463 food categories (each of which comprised multiple food items that are similar with respect to microbiological contamination likelihood) that had to be mapped onto over 65,000 food items reported by Nielsen. Because Nielsen did not provide the level of detail required to accommodate the food category breakdowns suggested by the experts we combined those subcategories that had a relatively low (≤ 2) weighted mean relative contamination likelihood score across all pathogens. As a result of this process, the total number of food categories reduced from 463 to 358.

Step 2 – Normalize weighted mean relative contamination likelihood scores such that the sum of the scores across food categories for a food pathogen equals 100 percent. The categories by which the experts indicated the likelihood of contamination associated with each food-pathogen combination in Round 2 will be converted to numerical values, ranging from 1 = Negligible to 8 = High:High (limitations of the assumption of ordinal ranking is discussed in Section 6). Next, we will aggregate the expert scores for each food-pathogen combination by calculating the numerical average over all experts and weighted each expert's score by his or her self-assessed confidence rating. This represents the weighted mean relative contamination likelihood score for that food product-pathogen combination. Then, for each pathogen, we will sum the weighted mean relative contamination likelihood score for each food category over all food categories linked to that pathogen. We will then normalized the weighted mean relative contamination likelihood score by calculating a weight for each food category as its weighted mean contamination likelihood score divided by the sum of all contamination likelihood scores for foods linked to that pathogen.

Step 3 – Use Nielsen sales equivalent units as proxy for consumption volume. As noted, another weighting factor must be introduced to account for each food category's share of consumption of the food categories relevant to each pathogen. Thus, this step accounted for relative consumption patterns. For each pathogen, we estimated the relative share of each food category in total consumption of all foods linked to that pathogen. As described earlier, we used secondary data rather than expert opinion to calculate relative consumption shares.

We mapped the Nielsen scanner data at the UPC level to the expert elicitation food category breakdowns using the available subcategory, product module, product, type, and container information. Best professional judgment was used to distribute total sales where available data lacked sufficient detail to match the expert elicitation food category breakdowns. In addition, some food category breakdowns originally requested by the experts, particularly those where relative contamination likelihood scores were not significantly different, were collapsed to a more aggregate category. This step facilitated a complex mapping without losing a great deal of granularity. Finally, serving size adjustments were made using best professional judgments of the analysts about the likely number of servings per sales equivalent unit for each food category.

Step 4 – Calculate raw attribution rate as the product of normalized weighted mean relative contamination likelihood score and consumption share in percentage terms. The percent of illnesses attributable to a specific food-pathogen combination will be calculated by multiplying the normalized weighted mean relative contamination likelihood score and the relative share of each food category in total consumption of all foods linked to that pathogen to get the raw attribution rate.

Step 5 – Normalize raw attribution rate such that the sum of the attribution rates for each food for a given pathogen equals 100 percent. The raw attribution rates will be re-normalized for each pathogen to ensure the model accounts for 100 percent of pathogen-specific illness; this will be the FBI attribution rate for each food category linked to that pathogen.

As an example, assume that only two food product categories were associated with source contamination from *Brucella* by the experts (see Table 2). Let us assume that the weighted mean relative contamination likelihood score for “Meat, Raw (not frozen)” is 2.29 and the ranking for “Milk and Cream – Unpasteurized” is 4.12, so the relative scores are $0.36 (= 2.29 \div [2.29 + 4.12])$ and $0.64 (= 4.12 \div [2.29 + 4.12])$, respectively. This leads to the relative likelihood rankings for all food product categories associated with *Brucella* to sum to 1.0, so they can be used to represent attribution fractions.

To extend the simple example in the case of *Brucella* (see Table 3), “Meat, Raw (not frozen)” accounts for 98 percent of consumption and “Milk and Cream – Unpasteurized” accounts for 2 percent of consumption. Therefore, among the food product categories that account for foodborne illnesses due to *Brucella*, 97 percent of illnesses are attributable to “Meat, Raw (not frozen),” and 3 percent to “Milk and Cream – Unpasteurized.” In other words, despite the higher relative risk of consuming unpasteurized

milk, in this illustrative example, fewer illnesses are attributed to it because only a small amount of unpasteurized milk is consumed by US consumers.

4 RESULTS

The expert elicitation process eliminated 9 percent (i.e., 182 out of 2,112) of original food-pathogen combinations judged to be very low risk for foodborne illness and identified new food sub-categories of concern for contamination. For some pathogens the experts considered a very small number of food-pathogen pairs to be of concern (see Table 4). For example, of the 96 original food subcategories, only two food categories, raw meat, unpasteurized milk and cream were considered to be relevant for *Brucella* and experts did not need to subdivide these categories to account for differences in the relative likelihoods of being contaminated as the product sits on the shelf. Conversely, experts subdivided the original 96 food subcategories into 353 relevant food subcategories to express differing levels of likelihood of contamination for *Salmonella*. Table 4 shows the number of relevant food subcategories experts identified for each pathogen.

The experts identified several food categories as commonly associated with high levels of contamination with more than one pathogen. For example, raw seafood was rated to have a higher likelihood of contamination with *Vibrio* spp., *Toxoplasma gondii*, Norwalk-like viruses, Rotavirus, Hepatitis A, *Listeria monocytogenes*, and Astrovirus. In contrast, condiments and sauces (except gravy) category was only rated at a higher likelihood of contamination with *Streptococcus*. The experts also indicated that *Salmonella*, *E. Coli*, *Campylobacter*, and *Listeria monocytogenes* pose a greater likelihood of contamination for foods (as indicated by their high relative likelihood of contamination scores) in comparison to other pathogens, such as Astrovirus, *Cyclospora cayetanensis*, *Giardia lamblia*, Rotavirus, *Toxoplasma gondii*, and *Trichinella spiralis* (as indicated by their low relative likelihood of contamination scores). The work to develop these inputs into attribution fractions will be the next step of this research.

From the second part of the elicitation, the results suggest that, according to these experts, the majority of foodborne illness cases for most pathogens are introduced to the product after the product leaves the store shelf. As mentioned earlier, this finding is similar to a finding in a recent 2011 study conducted by Batz et al. (2011) which finds that food handling and preparation problems in retail and food service settings account for a considerable burden of foodborne illnesses, even though their methodology was quite

different. However, for some pathogens, such as *Brucella*, *Vibrio* spp., *Cryptosporidium parvum*, *Cyclospora cayetanensis*, and *Giardia lamblia*, our expert elicitation findings suggest that the majority of foodborne illness cases are due to contamination that happens before the product reaches the store shelf.

5 DISCUSSION

5.1 Advantages of the Expert Elicitation Method

Expert elicitation provides a useful method when data are insufficient for meta-analysis or other types of systematic review to yield needed answers. For attributing foodborne illnesses to foods, experts in microbiology, food safety, food technology, and epidemiology have considerable relevant knowledge may not be available in an easily usable way in published literature. Expert elicitation also allows for synthesis of input from multiple types of experts, each of which may contribute unique insights into the problem.

The combination of expert elicitation with food consumption data allowed experts to focus on source contamination risk from highly disaggregated food categories. Combining the expert elicitation results with the food consumption data showed the food-pathogen combinations with high relative likelihoods of contamination may, nonetheless, result in lower numbers of foodborne illnesses than food-pathogen combinations with lower relative likelihoods of contamination, if the food with a high relative likelihood of contamination is eaten at low levels. For example, unpasteurized milk has a much higher relative likelihood of contamination with *Brucella* than raw meat, but because very little unpasteurized milk is consumed, few foodborne illnesses are associated with it in this study.

As a result of this approach, we were also able to address two significant concerns in designing an expert elicitation: the well-established “anchoring” and “availability” cognitive biases. Because most of the data available on attribution is based on outbreaks, it is likely that outbreak-based attribution will be the anchor from which experts adjust. They may try to adjust for the known limitations of outbreak data, but research shows that their adjustments are likely to be underestimates. The “availability” bias would lead experts to inadvertently overweight food categories for which outbreaks happened more recently. The approach we took of having the experts rank the relative likelihood of contamination as the product sits on the store shelf is less likely to be influenced by these biases. The experts were asked to calibrate their awareness of illnesses caused by each food commodity (from outbreak data) with their knowledge on the likelihood of

this food type being contaminated and the growth potential for this food type. By asking for a slightly different type of estimate, we may have been able to better overcome the impact of these cognitive biases.

There is a wide range of opinions in the literature about the validity of using self-assessed expertise scores to weight elicited responses. Because we knew there was heterogeneity in the levels of expertise by pathogen, we incorporated the self-rating. This was a short-cut to assessing expertise using other methods (Cooke, Soll, etc.), and may have impacted the results.

In assessing relative likelihood of contamination instead of directly assessing attribution fractions, we aimed to accomplish several things. First, because there were more different types of experts that had information about likelihood of contamination than those who had information about attribution fractions, we were able to develop a more robust estimate since it included expertise from a wider range of experts. Second, we diminished the influence of outbreak-based attribution fractions because of the known uncertainties around outbreak data. Finally, we added an element related to consumption to account for the fact that foods consumed at higher rates are more likely to cause illness, if the likelihood of contamination is the same.

5.2 Additional Considerations

Another consideration is the use of ordinal scales (high/medium/low/negligible) and the way that we interpreted that scale in the model. The experts were not given a numerical scale along with the verbal scale, so it is possible that they had different interpretation for this scale than was used in the model. For example, they may have been thinking that “high” was 3x as likely as low, and “medium” was 2x as likely as low. However, as described earlier, we used a linear scale. It would be valuable as part of a sensitivity analysis to see how much that scaling assumption impacts the results. Because the question focused on likelihood of contamination as the product sits on the shelf rather than likelihood of illness once the product is consumed, many of the confounding factors like storage, transport, personal and retail handling, cooking, and cross-contamination were eliminated. The experts did not have to make judgments that synthesized all those factors – they only had to focus on the product on the shelf and what happened to it before it got there. A simpler elicitation should lead to more robust results.

Validation of the model results are another significant challenge for EE. By the nature of the problem, there is little in the published literature to assess the validity of the results. To ensure the quality of our

results, we performed internal quality control reviews of the expert elicitation results to ensure that experts answered all questions as intended. We monitored inconsistencies and patterns of persistent bias by reviewing and comparing responses. Additionally, the elicitation protocol enabled participants to submit comments for each question, which we reviewed along with the question response to ensure that each question was interpreted as intended. We also plan to compare our final results with other studies.

6 CONCLUSIONS

This study addresses the challenges in finding the needed data for policy, in this case, understanding the relative contribution of different types of food to foodborne illness. Targeting resources most effectively to food safety activities depends on understanding the relative contribution of different foods to foodborne illnesses, but needed data is scarce and difficult to elicit. Expert elicitation provides a method to provide estimates, and for this situation, expert elicitation provides an approach for estimating model parameters and factors that are of interest. In cases such as those for this study, it may be the best, or only, way to obtain such estimates. Innovative approaches were developed to reduce the impact of cognitive biases and to reduce the cognitive burden on the experts.

WORKS CITED

- Batz, M., Hoffman, S., & Morris, J. (2011). *Ranking the Risks: The 10 Pathogen-food Combinations that have the Greatest Burden on Public Health*. University of Florida Emerging Pathogens Institute.
- Batz, M., Hoffman, S., Krupnick, A., Morris, J., Sherman, D., Taylor, M., et al. (2004). Identifying the Most Significant Microiological Foodborne Hazards to Public Health: A New Risk Ranking Model. *Food SAfty Research Consortium Discussion Paper Series, 1*, 1-26.
- Hoffmann, S., Fischbeck, P., Krupnick, A., & McWilliams, M. (2007). Using Expert Elicitation to Link Foodborne Illnesses in the United States to Foods. *Journal of Food Protection*, 70 (5), 1220-1229.
- Karns, S., Muth, M., & Cogliaiti, M. (2005, September). *Relative Risks of Meat and Poultry Products: An Expert Elicitation*. Retrieved from [http://www.fsis.usda.gov/PDF/Elicitation_Memo_092205 .pdf](http://www.fsis.usda.gov/PDF/Elicitation_Memo_092205.pdf)
- Painter, J. A., Hoekstra, R. M., Ayers, T., Tauxe, R. V., Braden, C. R., Angulo, F. J., et al. (2013). Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998–2008. *Emerging Infectious Diseases*, 19 (3).
- Pires, S., Evers, E., Pelt, W. v., Ayers, T., Scallan, E., Angulo, F., et al. (2009). Attributing the Human Disease Burden of Foodborne Infections to Specific Sources. *Foodborne Pathogens and Diseases*, 6 (4), 417-424.
- Scallan, E., Griffin, P., Angulo, F., Tauxe, R., & Hoekstra, R. (2011). Foodborne Illness Acquired in the United States - Unspecified Agents. *Emerging Infectious Diseases*.
- Scallan, E., Hoekstra, R., Angulo, F., Tauxe, R., Widdowson, M., Roy, S., et al. (2011). Foodborne Illness Acquired in the United States - Major Pathogens. *Emerging Infectious Diseases*.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.

Table 1: Composition of the Expert Panel

Current Professional Affiliation	Count	Percent of Total
Academia	5	31%
Government	4	25%
Industry/Private consultants	7	44%
Total	16	100%
Field of Expertise	Count	Percent of Total [a]
Microbiology	10	63%
Food Safety	6	38%
Epidemiology	3	19%
Food Technology	2	13%
Area of Expertise (Food Industry)	Count	Percent of Total [a]
Seafood	11	69%
Dairy	11	69%
Produce	10	63%
Meat & Poultry	10	63%
Shelf-stable Foods	10	63%
Canned Foods	13	81%
Refrigerated Foods	14	88%
Other	7	44%

[a] Percentages do not sum to 100% due to panel members with multiple areas of expertise

Table 2: Brucella – Example Attribution Rates Calculation – Steps 1 and 2

Food Product Category	Weighted Mean Relative Contamination Likelihood Score [a]	Normalized Weighted Mean Relative Contamination Likelihood Score		Comment
Meat, Raw (not frozen)	2.29	$2.29 \div 6.41$	0.36	
Milk and Cream – Unpasteurized	4.12	$4.12 \div 6.41$	0.64	<i>1.8 times likely to be contaminated than meat, raw (not frozen)</i>
Total	6.41		1.00	

[a] Weighted by expert's self-assigned confidence rating

Table 3: Brucella – Example Attribution Rates Calculation – Steps 3, 4, and 5

Food Product Category	Normalized Weighted Mean Relative Contamination Likelihood Score	Consumption Share	Raw Attribution Rate	Normalized Attribution Rate
Meat, Raw (not frozen)	0.30	98.31%	0.35	$0.35 \div 0.36 = 97\%$
Milk and Cream – Unpasteurized	0.54	1.69%	0.01	$0.01 \div 0.36 = 3\%$
Total	1.00	100%	0.36	100.00%

Table 4: Pathogens of Interest and Relevant Food Subcategories as Determined in Round 1

Pathogen	Number of Relevant Food Subcategories
<i>Bacillus cereus</i>	121
<i>C. botulinum</i>	110
<i>Brucella</i>	3
<i>Campylobacter</i>	45
<i>Clostridium perfringens</i>	67
<i>Escherichia coli</i> spp.	231
<i>Listeria monocytogenes</i>	172
<i>Salmonella</i> spp.	353
<i>Shigella</i>	116
<i>Staphylococcus</i>	96
<i>Streptococcus</i>	14
<i>Vibrio</i> spp.	35
<i>Yersinia enterocolitica</i>	32
<i>Cryptosporidium parvum</i>	102
<i>Cyclospora cayetanensis</i>	71
<i>Giardia lamblia</i>	31
<i>Toxoplasma gondii</i>	14
<i>Trichinella spiralis</i>	4
Norwalk-like viruses	135
Rotavirus	26
Astrovirus	14
Hepatitis A	138
Total	1,930