Foodborne illness source attribution estimates for 2019 for *Salmonella*, *Escherichia coli* O157, *Listeria monocytogenes*, and *Campylobacter* using multi-year outbreak surveillance data, United States

The Interagency Food Safety Analytics Collaboration (IFSAC)

October 2021
Acknowledgments

The methods used for this report were developed by members of the Interagency Food Safety Analytics Collaboration (IFSAC), which includes the Centers for Disease Control and Prevention (CDC), the U.S. Food and Drug Administration (FDA), and the U.S. Department of Agriculture’s Food Safety and Inspection Service (USDA-FSIS).

The following persons (listed alphabetically) contributed to or approved this report:

**CDC**
- Beau B. Bruce, MD, PhD
- Eleanor Click, MD, PhD
- LaTonia Clay Richardson, PhD, MS (*project lead*)
- Erica Billig Rose, PhD, MS

**FDA**
- Michael Batz, MSc
- Michael Bazaco, PhD, MS
- Cary Chen Parker, MPH
- Christopher Waldrop, MPH

**FSIS**
- Andrea Cote, DVM, MPH, DACVPM
- Neal Golden, PhD
- Sally Ann Iverson, DVM, MPH, DACVPM

**IFSAC Steering Committee**

**CDC**
- Patricia M. Griffin, MD
- Robert V. Tauxe, MD, MPH

**FDA**
- Katherine Vierk Brill, MPH
- Beverly Wolpert, PhD, MS

**FSIS**
- Kis Robertson Hale, DVM, MPH, DACVPM
- Joanna Zablotsky Kufel, PhD, MPH

**Suggested Citation**
**Executive Summary**

Each year in the United States an estimated 9 million people get sick, 56,000 are hospitalized and 1,300 die of foodborne disease caused by known pathogens. These estimates help us understand the scope of this public health problem. However, to develop effective prevention measures, we need to understand the types of foods contributing to the problem.

The Interagency Food Safety Analytics Collaboration (IFSAC) is a tri-agency group created by the Centers for Disease Control and Prevention (CDC), the U.S. Food and Drug Administration (FDA), and the U.S. Department of Agriculture’s Food Safety and Inspection Service (USDA-FSIS). IFSAC developed a method to estimate the percentages of foodborne illness attributed to certain sources using outbreak data from 1998 through the most recent year for four priority pathogens: *Salmonella*, *Escherichia coli* O157, *Listeria monocytogenes* and *Campylobacter*. IFSAC described this method and the estimates for 2012 in a report, a peer-reviewed journal article and at a public meeting.

IFSAC derived the estimates for 2019 using the same method used for the 2012 estimates, with some modifications. The data came from 1,532 foodborne disease outbreaks that occurred from 1998 through 2019 and for which each confirmed or suspected implicated food was assigned to a single food category. The method relies most heavily on the most recent five years of outbreak data (2015-2019). Foods are categorized using a scheme IFSAC created that classifies foods into 17 categories that closely align with the U.S. food regulatory agencies’ classification needs.

*Salmonella* illnesses came from a wide variety of foods.

More than 75% of *Salmonella* illnesses were attributed to seven food categories: Chicken, Fruits, Pork, Seeded Vegetables (such as tomatoes), Other Produce (such as nuts), Turkey and Eggs.

*E. coli* O157 illnesses were most often linked to Vegetable Row Crops (such as leafy greens) and Beef.

Over 75% of illnesses were linked to these two categories.

*Listeria monocytogenes* illnesses were most often linked to Dairy Products and Fruits.

More than 75% of illnesses were attributed to these two categories, but the rarity of *Listeria monocytogenes* outbreaks makes these estimates less reliable than those for other pathogens.

Non-Dairy *Campylobacter* illnesses were most often linked to Chicken.

Over 80% of non-Dairy foodborne illnesses were attributed to Chicken, Other Seafood (such as shellfish) and Turkey, with *Campylobacter* illnesses most often linked to Chicken. An attribution percentage for Dairy is not included because, among other reasons, most foodborne *Campylobacter* outbreaks were associated with unpasteurized milk, which is not widely consumed, and we think these over-represent Dairy as a source of illness caused by *Campylobacter*. Removing Dairy illnesses from the calculations highlights important sources of illness from widely consumed foods, such as Chicken.

This collaborative effort to provide annual attribution estimates continues IFSAC’s work to improve foodborne illness source attribution, which can help inform efforts to prioritize food safety initiatives, interventions and policies for reducing foodborne illnesses. These consensus estimates allow all three agencies to take a consistent approach to identifying food safety priorities to protect public health. For more information on IFSAC projects, visit https://www.cdc.gov/foodsafety/ifsac/projects/index.html.
Introduction
Each year in the United States, an estimated 9 million people get sick, 56,000 are hospitalized and 1,300 die of foodborne disease caused by known pathogens. These estimates help us understand the scope of this public health problem. However, to develop effective prevention measures, we need to understand the percentage of foodborne illnesses associated with specific foods; we call this work foodborne illness source attribution.

With the creation of the Interagency Food Safety Analytics Collaboration (IFSAC) in 2011, the Centers for Disease Control and Prevention (CDC), the U.S. Food and Drug Administration (FDA), and the U.S. Department of Agriculture’s Food Safety and Inspection Service (USDA-FSIS) agreed to improve data and methods used to estimate foodborne illness source attribution and provide timely estimates of the food sources of four priority foodborne pathogens: *Salmonella*, *Escherichia coli* O157, *Listeria monocytogenes* and *Campylobacter*. In this report, we use the term *Listeria* to refer to *Listeria monocytogenes*. IFSAC considers these priority pathogens because of the frequency (estimated 1.9 million illnesses each year combined) and severity of illness they cause, and because targeted interventions can significantly reduce these illnesses.

IFSAC developed a method for analyzing outbreak data to estimate which foods are responsible for illnesses related to the four priority pathogens, using a scheme IFSAC created to classify foods into 17 categories that closely align with the U.S. food regulatory agencies’ classification needs. IFSAC described this method and the resulting estimates for the year 2012 in a report, a peer-reviewed article, and at a public meeting. IFSAC derived the estimates for 2019 using the same method, with some modifications. IFSAC publishes annual estimates of the sources of foodborne illness for the priority pathogens while continuing to work on methods to further improve these estimates.

Consensus among the three agencies on methods and attribution estimates can help inform efforts to prioritize food safety initiatives, interventions and policies for reducing foodborne illnesses. The 2019 estimates achieve IFSAC’s goals of using improved methods to develop estimates of foodborne illness source attribution for priority pathogens and of achieving consensus that these are the best current estimates for the agencies to use in their food safety activities. These estimates can also help scientists; federal, state and local policy-makers; the food industry; consumer advocacy groups; and the public to assess whether prevention measures are working.
Methods
We analyzed data extracted on March 4, 2021 from CDC’s Foodborne Disease Outbreak Surveillance System (FDOSS)5,6 (www.cdc.gov/foodsafety/fdoss) on finalized reports of outbreaks confirmed or suspected to be caused by the four priority pathogens in which the first illness onset occurred from 1998 through 2019. We excluded outbreaks that met one or more of the following conditions: occurred in a U.S. territory; had no identified food vehicle or contaminated ingredient; were caused by more than one pathogen (including pathogens not included in this report). We excluded outbreaks that were caused by both Salmonella serotype Enteritidis and another Salmonella serotype because we modeled Enteritidis and non-Enteritidis outbreaks separately. Similarly, because this analysis does not include non-O157 STECs, we excluded outbreaks caused by both E. coli O157 and another E. coli serogroup.3 We included outbreaks caused by multiple serotypes of Salmonella if none were Enteritidis, and outbreaks caused by multiple species of Campylobacter.

Each outbreak was assigned to a single food category using the IFSAC food categorization scheme2 based on confirmed or suspected implicated foods and ingredients (i.e., a single ingredient was confirmed or suspected to be implicated or all ingredients in the food were assigned to the same food category). We excluded outbreaks that could not be assigned to a single food category, usually because the food was complex (i.e., composed of ingredients belonging to more than one category) and the contaminated ingredient in the complex food could not be identified.

We developed pathogen-specific analysis of variance models using our previously described method3 to mitigate the impact of large outbreaks and control for epidemiological factors. We estimated the number of log-transformed illnesses associated with each outbreak based on three factors deemed to be important based on our exploratory analyses: food category, type of preparation location (e.g., restaurant, home) and whether the outbreak occurred in one or more states. We modeled outbreaks caused by Salmonella Enteritidis separately from other serotypes because we found Enteritidis outbreaks to have meaningful differences in outbreak size and epidemiological factors. We summed results to obtain combined Salmonella estimates.

These model estimates were then back-transformed and down-weighted with a function that declines exponentially for outbreaks older than the most recent five years (2015-2019) because we considered foods more recently implicated to be most relevant for estimating current attribution.

We used the resulting down-weighted model-estimated illnesses to calculate each estimated attribution percentage: the sum of illnesses associated with a pathogen-food category pair was divided by the sum of illnesses associated with that pathogen across all food categories. We calculated 90% credibility intervals and considered non-overlapping credibility intervals an indication of statistical significance at the p<0.10 level. After down-weighting, 71% of overall information came from the most recent five years, 25% from the next most recent five years (2010-2014) and 5% from the oldest data (1998-2009).

In the graphs and tables, food categories appear in descending order of their estimated attribution percentage, and those that contributed to a cumulative attribution of approximately 75% of illnesses are indicated.
Results
We identified 4,183 outbreaks that occurred from 1998 through 2019 and that were confirmed or suspected to be caused by *Salmonella*, *E. coli* O157, *Listeria* or *Campylobacter*, including 198 outbreaks that were confirmed or suspected to be caused by multiple pathogens or serotypes. Of these, we excluded 89 outbreaks according to our pathogen exclusion criteria, leaving 4,094 outbreaks. We further excluded 1,665 outbreaks without a confirmed or suspected implicated food, 893 outbreaks for which the food vehicle could not be assigned to one of the 17 food categories and four that occurred in a U.S. territory.

The resulting dataset included 1,532 outbreaks in which the confirmed or suspected implicated food or foods could be assigned to a single food category: 943 caused or suspected to be caused by *Salmonella*, 269 by *E. coli* O157, 51 by *Listeria* and 269 by *Campylobacter.* These include 51 outbreaks caused by multiple serotypes of *Salmonella* (43) and multiple species of *Campylobacter* (8). Due to down-weighting, the last five years of outbreaks provide the majority of information for the estimates; outbreaks from 2015 through 2019 provide 73% of model-estimated illnesses used to calculate attribution for *Salmonella*, 65% for *E. coli* O157, 62% for *Listeria* and 56% for *Campylobacter*. The overall results and those for each pathogen are shown in Figures 1 through 5.

*Among the 269 *Campylobacter* outbreaks that occurred during the study period, the 155 assigned to Dairy were excluded from the final calculation of the attribution percentage. Thus, the *Campylobacter* attribution point estimates reflect data from 114 outbreaks.*
Overall Key Results

- The results are based on 943 outbreaks caused or suspected to be caused by *Salmonella*, 269 by *E. coli* O157, 51 by *Listeria* and 114 by *Campylobacter* (after 155 outbreaks due to Dairy were excluded).
- Estimated *Salmonella* illnesses were more evenly distributed across food categories than illnesses from *Campylobacter, E. coli O157* and *Listeria*; most of the illnesses for the latter pathogens were attributed to one or two food categories.
- The credibility intervals overlap for the *Salmonella* and *Listeria* categories with the highest attribution percentages, indicating no statistically significant difference among them.
Figure 2: Estimated percentage of foodborne *Salmonella* illnesses (with 90% credibility intervals) for 2019, in descending order, attributed to each of 17 food categories, based on multi-year outbreak data,* United States. Click here to download relevant data.

*Based on a model using outbreak data that gives equal weight to each of the most recent five years of data (2015-2019) and exponentially less weight to each earlier year (1998-2014).

**Salmonella Key Results**

- Over 75% of illnesses were attributed to seven food categories: Chicken, Fruits, Pork, Seeded Vegetables (such as tomatoes), Other Produce (such as fungi, herbs, nuts and root vegetables), Turkey and Eggs.
- The credibility intervals for each of the seven food categories that account for 75.9% of all illnesses overlap with some of the others.
Figure 3: Estimated percentage of foodborne *Escherichia coli* O157 illnesses (with 90% credibility intervals) for 2019, in descending order, attributed to each of 17 food categories, based on multi-year outbreak data,* United States. Click here to download relevant data.

*Based on a model using outbreak data that gives equal weight to each of the most recent five years of data (2015-2019) and exponentially less weight to each earlier year (1998-2014).

**E. coli O157 Key Results**

- Over 75% of *E. coli* O157 illnesses were attributed to Vegetable Row Crops (such as leafy greens) and Beef.
- Vegetable Row Crops had a significantly higher estimated attribution percentage than all other categories.
- Beef had a significantly higher estimated attribution percentage than all categories other than Vegetable Row Crops.
- No illnesses were attributed to Eggs or Oils-Sugars.
**Figure 4:** Estimated percentage of foodborne *Listeria monocytogenes* illnesses (with 90% credibility intervals) for 2019, in descending order, attributed to each of 17 food categories, based on multi-year outbreak data,* United States. Click here to download relevant data.

*Based on a model using outbreak data that gives equal weight to each of the most recent five years of data (2015-2019) and exponentially less weight to each earlier year (1998-2014).

**Listeria monocytogenes Key Results**
- Over 90% of illnesses were attributed to non-meat food categories.
- Over half of illnesses were attributed to Dairy and Fruits.
- Among our annual reports, which began for the year 2012, this is the first in which *Listeria* illnesses were attributed to either Other Produce or to Eggs. The Other Produce attribution is influenced by a multistate outbreak due to enoki mushrooms detected in 2019, with cases during 2016-2019. The Egg attribution comes from a multistate outbreak due to hard-boiled eggs detected in 2019, with cases during 2017-2019.
- The credibility intervals for the Dairy, Fruits, Vegetable Row Crops and Other Produce categories were quite wide, partly due to the small total number of outbreaks (51). The credibility intervals overlapped each other, and the intervals for the Fruits, Vegetable Row Crops and Other Produce categories overlapped those for some food categories with much smaller estimated attribution percentages, such as Sprouts and Eggs.
- No illnesses were attributed to Other Meat/Poultry, Game, Other Seafood, Grains-Beans, Oils-Sugars or Seeded Vegetables.
Figure 5: Estimated percentage of foodborne *Campylobacter* illnesses (with 90% credibility intervals) for 2019, in descending order, attributed to each of 16 food categories (dairy was excluded), based on multi-year outbreak data,*† United States. Click here to download relevant data.

*Based on a model using outbreak data that gives equal weight to each of the most recent five years of data (2015-2019) and exponentially less weight to each earlier year (1998-2014).
† *Campylobacter* estimates exclude results derived from Dairy outbreak data.

**Campylobacter Key Results**
- Nearly two-thirds (64.7%) of non-Dairy *Campylobacter* illnesses were attributed to Chicken, significantly higher than any other food category.
- The only other food categories with attribution percentages over 5% were Other Seafood, Turkey and Other Meat/Poultry.
- The majority (36/64, 56%) of chicken-associated *Campylobacter* outbreaks were attributed to chicken liver products, which are not widely consumed in the United States.
- No illnesses were attributable to Eggs, Grains-Beans or Sprouts.
- An attribution percentage for Dairy is not presented partly because most Dairy-related foodborne *Campylobacter* outbreaks were associated with unpasteurized milk (139/155, 90%), which is not widely consumed in the United States. The attribution percentages before removing Dairy were Dairy 51.9%, Chicken 31.1%, Other Seafood 4.9%, Turkey 3.2% and Other Meat/Poultry 2.8%; they were less than 2% for other categories.
Discussion

This report uses data from 1998 through 2019 to provide outbreak-based attribution estimates for 2019 of the percentage of illnesses caused by four priority pathogens, attributing illnesses to each of 17 food categories. Data from foodborne disease outbreaks are the foundation of many foodborne illness source attribution analyses, in part because outbreak investigations often link illnesses to a specific food and the data are captured nationally. An IFSAC study using data from the years 2004 to 2011 found that outbreak and sporadic infections caused by the four priority pathogens were generally demographically similar; this finding supports the use of foodborne outbreaks for source attribution. These estimates can inform food safety decision making and provide pathogen-specific and product-specific direction for reducing foodborne illness.

The attribution of *Salmonella* illnesses to multiple food categories suggests that interventions designed to reduce illnesses from *Salmonella* need to target a variety of food categories. In contrast, the majority of *E. coli* O157 illnesses were attributed to two food categories: Vegetable Row Crops and Beef. These data suggest that interventions for *E. coli* O157 focusing on these two food categories may be most effective in reducing illnesses.

As with *E. coli* O157, most *Listeria* illnesses were also attributed to two food categories: Dairy and Fruits. Although the limited number of outbreaks and wide credibility intervals dictate caution in interpreting the attribution percentage for Dairy, the risk to pregnant women and persons with weakened immune systems of consuming soft cheese made from unpasteurized milk or in unsanitary conditions is well-recognized, and outbreaks from Fruits have been observed in recent years.

Our method attributed most *Campylobacter* illnesses to Chicken. The majority of chicken outbreaks used in the analysis were due to chicken liver products, which are not widely consumed. An analysis of (National Health and Nutrition Examination Survey) NHANES data found that only about 0.1% of all chicken products consumed are liver or liver-derived. Further, the attribution to Chicken is high because the attribution percentages for Dairy are not presented in the figures for *Campylobacter*. Dairy outbreaks were not included in this analysis because most of the reported *Campylobacter* Dairy outbreaks were associated with unpasteurized milk (139/155, 90%), which is not widely consumed by the general population. Moreover, an analysis of 38 case-control studies of sporadic campylobacteriosis found a much smaller percentage of illnesses attributable to consumption of raw milk than chicken. For example, a U.S. FoodNet case-control study attributed 1.5% of campylobacteriosis cases to consumption of unpasteurized milk, compared with 24% to consumption of chicken prepared in a restaurant. Structured expert judgment studies estimate about 8-10% of foodborne campylobacteriosis to be attributable to dairy products (principally raw milk), compared with 33-72% to chicken. Thus, *Campylobacter* outbreaks in the Dairy food category appear to over-represent Dairy as a source of *Campylobacter* illness. After removing the Dairy outbreaks, the Chicken attribution increased to 64.7%, which is consistent with published literature. The chicken liver and raw dairy consumption data indicate limitations in our method for estimating attribution fractions for *Campylobacter*.

This analysis includes eight outbreaks (7 *Listeria* and 1 *Salmonella*) that spanned multiple years and were not included in prior reports because the investigations had not been completed. Each of the outbreaks had a first

---

†FSIS analysts examined nine two-year cycles (1996-2016) of National Health and Nutrition Examination Survey (NHANES) data constituting 18 years of food consumption data. They estimated the U.S. population average daily consumption of chicken livers and all chicken products using population weights provided.
illness onset date before 2019 and a last illness onset date during or after 2019. These outbreaks were solved in large part because whole-genome sequencing linked human and food isolates from multiple years.

Our approach addresses several issues with outbreak-based foodborne illness source attribution, yet limitations associated with generalizing outbreak data to sporadic illnesses remain and are well-documented.5,6 Our analysis is also subject to other uncertainties and biases. First, for pathogens with a small number of outbreaks, outbreaks with a very large illness count can have substantial influence on the attribution point estimate. Second, this analysis only included 37% (1,532 of 4,183) of reported outbreaks caused by the four priority pathogens because we excluded those that did not meet our pathogen inclusion criteria, those in which the implicated food could not be assigned to a single food category and those that occurred in a U.S. territory. These outbreaks might not be representative of all outbreaks caused by these pathogens. Third, our analysis includes illnesses that occurred among institutionalized populations, such as those in prisons, hospitals and schools; these populations are easier to identify and collect complete data from, have fewer food options and are not representative of the general population. Finally, data from outbreaks are typically reported to CDC and verified during the year after the investigation ends. Due to the limited capacity of some states to report 2019 outbreaks during 2020 because of the coronavirus disease 2019 (COVID-19) pandemic, data from some jurisdictions may be incomplete, and this may have contributed to the uncertainty of this year’s estimates.

These estimates should not be interpreted as suggesting that all foods in a category are equally likely to transmit pathogens. Caution should also be exercised when comparing estimates across years, as a decrease in a percentage may result, not from a decrease in the number of illnesses attributed to that food, but from an increase in illnesses attributed to another food. This is especially true for Listeria, as the attribution percentages might vary widely from year to year due to the limited number of outbreaks and the zero-sum nature of the attribution percentages. The analyses show relative changes in percentage, not absolute changes in attribution to a specific food. Therefore, we advise using these results in conjunction with other scientific data for decision making.

Conclusions
IFSAC’s work to provide a harmonized analytic approach for estimating foodborne illness source attribution from outbreak data can provide consistency in the use and interpretation of estimates across public health and regulatory agencies. As more data become available and methods evolve, attribution estimates will improve. Annual updates to these estimates will enhance IFSAC’s efforts to inform and engage stakeholders, and further their ability to assess whether prevention-oriented measures are working.

IFSAC continues to enhance attribution efforts through projects that address limitations identified in this report. For more information on IFSAC’s completed and ongoing projects, visit http://www.cdc.gov/foodsafety/ifsac/index.html.


