







## General Interest

# An Overview of Traceback Investigations and Three Case Studies of Recent Outbreaks of *Escherichia coli* O157:H7 Infections Linked to Romaine Lettuce

KARI IRVIN<sup>1</sup>  <https://orcid.org/0000-0002-8942-4314>,<sup>1\*</sup> STELIOS VIAZIS<sup>2</sup>  <https://orcid.org/0000-0002-4711-1520>,<sup>1</sup>  
ANGELA FIELDS<sup>2</sup>  <https://orcid.org/0000-0002-4384-3417>,<sup>1</sup> SHARON SEELMAN,<sup>1</sup> KAREN BLICKENSTAFF,<sup>1</sup> ELLEN GEE,<sup>1</sup>  
MATTHEW E. WISE<sup>2</sup>  <https://orcid.org/0000-0001-8212-3458>,<sup>2</sup> KATHERINE E. MARSHALL<sup>2</sup>  <https://orcid.org/0000-0001-7220-7316>,<sup>2</sup>  
LAURA GIERALTOWSKI<sup>2</sup>  <https://orcid.org/0000-0002-0575-4579>,<sup>2</sup> AND STIC HARRIS<sup>1</sup>

<sup>1</sup>Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, 4300 River Road, College Park, Maryland 20740 (ORCID: <https://orcid.org/0000-0001-8788-1485> [S.S.]); and <sup>2</sup>Centers for Disease Control and Prevention, Atlanta, Georgia 30333, USA

MS 21-112: Received 16 March 2021/Accepted 5 April 2021/Published Online 9 April 2021

## ABSTRACT

Leafy greens contaminated with Shiga toxin–producing *Escherichia coli* have continued to cause foodborne illness outbreaks in recent years and present a threat to public health. An important component of foodborne illness outbreak investigations is determining the source of the outbreak vehicle through traceback investigations. The U.S. Food and Drug Administration is home to traceback investigation experts who use a standardized process to initiate, execute, and interpret the results of traceback investigations in collaboration with the Centers for Disease Control and Prevention and state and local partners. Traceback investigations of three outbreaks of Shiga toxin–producing *E. coli* infections linked to romaine lettuce in 2018 and 2019 were examined to demonstrate challenges, limitations, and opportunities for improvement. The three outbreaks resulted in a total of 474 illnesses, 215 hospitalizations, and 5 deaths. These illnesses were linked to the consumption of romaine lettuce from three distinct growing regions in Arizona and California. Some of the challenges encountered included the time it took to initiate a traceback, limited product-identifying information throughout the supply chain, lack of interoperability in record-keeping systems, and comingling of product from multiple suppliers. These challenges led to time delays in the identification of the farm source of the leafy greens and the inability to identify the root cause of contamination. Implementation of technology-enabled traceability systems, testing of these systems, and future regulations to incentivize adoption of traceability systems are some of the initiatives that will help address these challenges by improving traceback investigations and ultimately preventing foodborne illnesses and future outbreaks from occurring.

## HIGHLIGHTS

- Traceback investigations are a critical tool in the outbreak investigation process.
- Traceback methods for *E. coli* O157:H7 outbreaks linked to lettuce are discussed.
- Results of each traceback investigation led to the recovery of the outbreak strain.
- There are many challenges and limitations inherent in current traceback methods.
- Efforts are underway to address challenges and improve traceback investigations.

Key words: Foodborne illness outbreaks; Leafy greens; Shiga toxin–producing *Escherichia coli*; Traceback investigations

In the United States, vegetable row crops, which include leafy greens, are a frequent source of Shiga toxin–producing *Escherichia coli* (STEC) O157:H7 outbreak-related illnesses (10, 12, 13). Leafy greens contaminated with STEC are a significant public health problem and were the source of 40 outbreaks in the United States and Canada from 2009 to 2018 (18, 30, 32). Several factors may contribute to the contamination of leafy greens. These include, but are not limited to, contamination at the farm

level from the spread of animal feces via farm animals, wildlife, or rainwater runoff; contamination of irrigation water; and cross-contamination during harvesting and at the processing level that may occur when a batch of product from a contaminated field on a farm is comingled with other batches of product (4, 16). As a result, contamination at one farm can cross-contaminate product from other farms when combined to make mixed salad products. This leads to complex outbreak investigations as there is an increased likelihood of multiple contaminated products being available for consumption. Given the well-documented severity of STEC infections (14), the potential for clinical

\* Author for correspondence. Tel: 240-402-2946; Fax: 301-436-3221; E-mail: kari.irvin@fda.hhs.gov.

complications, and the fact that most leafy greens are commonly consumed raw, prevention strategies are important to minimize risk of contamination (39). However, once a foodborne illness outbreak does occur, rapid determination of the vehicle, its removal from the market, and identification of the contamination source are necessary to stop the ongoing outbreak and prevent future outbreaks. An important tool applied to achieve this goal is the traceback investigation. A traceback investigation is the process of reviewing product supply chain data to identify the source of food served or sold at a specific location, known as a point of sale (POS).

Key government agencies in the traceback investigation process reside at the local, state, and federal level. The U.S. Food and Drug Administration (FDA) is dedicated to investigating multistate foodborne illness outbreaks and serving as a traceback investigation expert (37). The Centers for Disease Control and Prevention (CDC), state, and local public health and regulatory partners routinely collaborate with the FDA on traceback investigations as part of multistate foodborne illness outbreak investigations (8). The evidence produced by traceback investigations is used to complement epidemiologic and laboratory evidence collected during foodborne illness outbreak investigations and can help inform the use of appropriate public health and regulatory tools aimed at removing the contaminated food from commerce and protecting public health. A traceback investigation may be conducted for one, or more, of the following reasons: (i) to identify the source(s) and distribution of the suspected or confirmed food product and remove it from commerce, (ii) to distinguish between two or more suspected food products, and (iii) to assist in determining the potential route(s) and/or source(s) of contamination and prevent future illnesses and outbreaks. Some of the intricacies and challenges of traceback investigations in the United States have been previously presented within the context of guidance for public health investigators (8) and briefly described during various foodborne illness outbreak investigations through the years (17, 19, 20, 25).

Traceback investigations are a critical component of any foodborne illness outbreak investigation and are especially important during outbreaks linked to produce, such as leafy greens, given the need to rapidly identify the product source(s), often without the availability of product labeling. Leafy greens are a challenging commodity for traceback investigations partly because they are widely and frequently consumed, typically eaten raw, and have a short shelf life (18). Leafy greens may be sold loose with no packaging or eaten at restaurants, situations that make it difficult for the consumer to identify the variety or brand. Furthermore, prepackaged green salad mixes may be sold under multiple brands and labels and include multiple leafy green products from numerous farms grown on separate ranches and fields. Contamination at any one of those farms, ranches, or processors can result in wide distribution of contaminated product in salad mixes. When traceback investigations identify a single source for a product or one of its ingredients, regulatory and public health actions can take place. In addition, work may be done to investigate the

root cause of the contamination and implement appropriate corrective actions. Here, we present the basic concepts and processes used to initiate and execute traceback investigations, interpret the results used to inform public health and regulatory action, and discuss the limitations and challenges that characterize them through the analysis of three recent outbreaks of STEC infections linked to leafy greens.

## TRACEBACK INVESTIGATIONS

**Background.** Traceback investigations are used to determine and document the movement of food product(s) through the supply chain by following the food product back through the points of distribution, processing, and production to determine the source of the product and/or its ingredients. A traceback investigation involves many steps. These include the following: analyzing epidemiologic information to select appropriate food exposures from ill persons to initiate traceback, contacting each point in the supply chain to obtain the correct available records, reviewing all available document records, and timely reporting of the results to investigational partners for regulatory decision making. To understand how the traceback process works, it is important to become familiar with the use of certain terminology. The terms used herein to describe the traceback process are defined in Table 1. These definitions are intended to help understand the process and terminology used during traceback investigations; however, additional definitions may be available in other contexts, such as guidelines, regulations, or other resources.

**Epidemiologic investigation.** Possible foodborne illness outbreaks can be detected through a variety of mechanisms, but in multistate foodborne illness outbreaks, they are most often detected by the CDC through PulseNet, the national molecular subtyping network for foodborne disease surveillance (5). After an outbreak is detected, investigators in local and state health agencies (and the CDC during multistate investigations) collect and/or analyze food histories through interviews of ill people to identify any common food exposures as well as POS locations from which traceback could be initiated. Information about the POS where food was purchased or served also assists in identification of illness subclusters. Local and state health officials work to obtain specific information about when and where ill people purchased or ate the food item of interest and determine whether documentation of the exposure is available, such as shopper cards, loyalty cards, and receipts. A shopper or loyalty card is a card issued by a grocery store (supermarket) or chain restaurant, respectively, to a customer and is used to capture consumer purchase data. As a result, the purchase data from these cards are very helpful in epidemiologic investigations. Lastly, local and state officials may also conduct investigations to determine the potential for cross-contamination at the POS.

**Selecting traceback illness subclusters.** Determining whether illness subclusters are suitable for traceback marks the starting point for a traceback investigation; it is one of

TABLE 1. *Definitions of terms used to describe the traceback process*

Term	Definition
Single case exposure	An ill person within a foodborne illness outbreak reporting a one-time consumption of a specific food before their illness that matches the known timeframe for illness onset based on the pathogen.
Foodborne illness outbreak	An incident in which two or more persons experience a similar illness caused by the same pathogen resulting from the ingestion of a common food.
Illness subcluster	A group of unrelated ill people within a foodborne illness outbreak who reported eating or purchasing food from the same point of sale (POS), or group of POS locations (e.g., same restaurant, different locations) before their illness began.
POS	The location where an illness subcluster or single ill person reported purchasing or consuming the food being traced (e.g., restaurants, grocery stores). For purposes of this article, this also includes facilities such as nursing homes and prisons.
Distribution center	A facility that stores and subsequently ships product from one location to the next without changing the product in any way.
Processor	A facility that receives ingredients and changes or combines those ingredients to create a new product. Herein, processors are those facilities that receive leafy greens and wash, chop, mix, and package the leafy greens into a variety of products.
Farm	Land devoted to the growing of crops, in this case leafy greens.
Grower	The person or business responsible for the growing of leafy greens at the farm.
Ranch or field	Separate subdivisions of the farm used to grow leafy greens. These terms may be used interchangeably depending on the regionality of the location. Ranch is commonly used in Arizona and California. Field is used more widely.
Records	Documents that confirm the movement of food from one place to another along the supply chain. Records consist of a wide variety of types, including invoices, purchase orders, bills of lading, and production logs. Records can be stored in paper or electronic format such as spreadsheets.
Traceback	Process of reviewing product supply chain records to identify the origin of food served or sold at a specific POS.
Traceback leg	The supply chain for a specific illness subcluster or single case exposure, initiated at the POS.
Supply chain	The path a food takes from origin to POS. Specific points along the supply chain include, but are not limited to, grocery stores, restaurants, distribution centers, processors, farms, and ranches.
Convergence	A point in the supply chain where multiple traceback legs share a commonality. Commonalities can include shared facilities, land, water sources, and geographic regions.
Timeframe of interest	The time period when contaminated product could have moved through the supply chain. This is often calculated by considering both the shelf life of the product and the product turnover rate at various points along the supply chain.

the most critical steps of the process and can affect the outcome of the investigation (Table 1). The FDA, with input from CDC, state, and local officials, determines which illness subclusters will be selected to begin the federal traceback. Typically, multiple illness subclusters are selected with the goal of representing both the geographic diversity of outbreak-associated illnesses and the time span of the outbreak period. The following criteria are considered when selecting illness subclusters for traceback initiation: (i) the bacteria cultured from the ill person is genetically closely related by whole genome sequencing to the outbreak strain (i.e., not a variant strain); (ii) a reliable food history (or histories) is available for the ill person; (iii) the ill person reported few, or preferably one, exposure to the suspect food before illness onset, which helps ensure the correct food is pursued and the proper exposure is considered; and (iv) verifiable purchase dates can be obtained, through receipts, loyalty cards, shopper cards, or other forms of documentation. The more ill people associated with an illness subcluster, the stronger the evidence that the contaminated product was sold at that specific POS. In practice, however, illness subclusters identified during outbreaks typically do not meet all the criteria noted above. In these instances, in an effort to move

quickly and determine the source of the suspected food, illness subclusters with the best available information are used for traceback initiation. As noted elsewhere, the more illness subclusters that are included in a traceback investigation, the greater the level of confidence when a common supplier is identified (18).

Illness subclusters identified early in the traceback investigation serve as starting points for traceback investigations, and additional illness subclusters may be added throughout the investigation. In the absence of illness subclusters or in situations where additional traceback legs are needed, single case exposures may be considered for inclusion in the traceback investigation, especially if the ill person has a single exposure to the suspect food and can provide strong documentation for their food purchase or exposure. Food exposure data may not always be available for every ill person; therefore, not all persons within the overall foodborne illness outbreak cluster are included in the traceback.

**Documentation collection and review.** Records from POS associated with illness subclusters and single case exposures that are included in the federal traceback are collected and thoroughly reviewed to determine the supply

chain for the suspect food. Record collection is conducted by local, state, and federal officials. Local and state health officials may initiate a traceback by collecting records at the POS. Working collaboratively, when a POS is included in the federal traceback investigation, state and local officials share the previously collected information with the FDA. This allows the FDA to continue the traceback investigations for firms that may fall outside of the jurisdiction of these local or state officials. In addition, this allows the FDA to tie together information from multiple jurisdictions and begin the process of identifying convergence for the suspected food.

The FDA conducts record collection through field assignments or information requests generated by the FDA Coordinated Outbreak Response and Evaluation Network, collaboratively with the appropriate subject matter experts, which are submitted to field offices in the FDA Office of Regulatory Affairs or the Office of Regulatory Affairs Produce Safety Network.

Based on actual or estimated dates of exposure to, or purchase of, the suspect food, a timeframe of interest is established to scope the record collection. If records are available, they are provided to government officials in the form of paper records or pulled from electronic databases in the form of spreadsheets. Regulatory officials seek information on stock rotation or shortage, delivery frequency, and shelf life of product (Table 2). Extensive investigations may take place along the supply chain to obtain information on each firm's receipt, handling, and/or processing of the food of interest by identifying the ordering, shipping and receiving, stock rotation, inventory, and preparation procedures while obtaining the appropriate records and documentation.

**Data analysis.** The FDA's Coordinated Outbreak Response and Evaluation Network conducts an analysis of the data available from the records and begins creating two documents that illustrate the traceback: a traceback timeline (Fig. 1) and a traceback diagram (Fig. 2). The timeline is a visual reference that illustrates the dates that correspond with each shipment of a food product during the specified timeframe of interest, at each level of distribution. Data elements may include date of shipment and receipt, quantity of product, lot code, batch code, addresses for shippers and receivers, product description, and any grower, harvester, and cooler information for the timeframe of interest. This information allows investigators to determine the shipments that most likely contained the contaminated product. In the example illustrated in Figures 1 and 2, the timeline shows that on 18 May, an exposure to chopped romaine lettuce occurred at Steak House. Incoming shipment data and the frequency of shipments available at Steak House received from two distribution centers, XYZ Produce Distributors and Nations Foods Distribution, can be visualized on the timeline. This visualization, along with information from the investigation about inventory management and stock rotation practices at Steak House, helps identify three shipments of interest from XYZ Produce Distributors that could have supplied the contaminated product to Steak

House. These shipments are implicated for further investigation; information from records provided by XYZ Produce Distributors is also captured. This process continues until the origin of the food for each implicated shipment is identified. In this example, shipments of romaine lettuce from Zenith Fresh Processing at XYZ Produce Distributors are determined to be the shipments that could have supplied the contaminated product in the implicated shipments to the restaurant. Investigation of the implicated shipments during the timeframe of interest identifies that product at Zenith Fresh Processing was provided by a single farm, World Farms, with sources from three ranches (ATP, ZPT, and MO).

Once these shipments of interest are identified, a traceback timeline and diagram are generated for each traceback leg to demonstrate the movement of the potentially contaminated product along the supply chain, from the POS back to its original source(s). Figure 1 shows the traceback timeline depicting the Steak House traceback leg. Figure 2 shows the traceback diagram constructed from the traceback timeline to better visualize the product supply chain within this specific traceback leg. A master traceback diagram is generated by combining the traceback diagrams for each traceback leg to illustrate convergence across all traceback legs traced within the overall outbreak (Figs. 3 through 6).

The completion of both the traceback timeline and diagrams can only be accomplished after the thorough review of data supplied during the record collection process. The records necessary for the food of interest often vary widely in terms of document types and formats and levels of completeness. Records collected may or may not be able to link an outgoing shipment from one firm to an incoming shipment to another or link what was received at a firm to what was shipped out of the same firm. Given the variability in records received, subsequent follow-up with a firm often occurs after the initial record collection to obtain clarity on the information provided. In the example noted above, neither Steak House nor XYZ Produce Distributors can determine which specific lot codes of product were available at Steak House on 18 May, so an assessment is made based on timing of the shipment, inventory control, production documentation, and stock rotation practices. This assessment based on timing expands the scope of the traceback, requiring three shipments from Zenith Fresh Processing to be investigated, versus potentially one or two. Although confidence in the traceback leg is not necessarily compromised based on these assessments, the lack of precision creates delays in the traceback process, as additional shipments must be investigated.

#### **Traceback investigation outcomes and significance.**

Traceback investigations are meant to identify potential convergence and are one way to identify a common source that can explain the food exposures reported by ill people in the illness subclusters or single case exposures selected for traceback. Epidemiologic analysis of food exposure data in foodborne illness outbreak investigations often generates a hypothesis that requires confirmation. Traceback investiga-



TABLE 2. *Types of records, information, and samples that were requested and collected during the three traceback investigations through assignments*

Information type	Type of records or information requested
Record collection	<p>Provide records for outgoing shipments of romaine lettuce from firm during the timeframe of interest, including the intended recipients. Share any records or spreadsheets if available.</p> <p>Provide explanations of brands, lots, codes, dates, or other identifiable information used to determine traceability of romaine lettuce.</p> <p>Provide records for incoming shipments to firm during the timeframe of interest.</p> <p>Determine whether there were changes in supplier or locally sourced products during the timeframe of interest.</p> <p>Is the firm able to link incoming shipments to outgoing shipments? Please describe the tracking process (i.e., traceability)? Indicate how the dates on the shipping records reflect the date the products were received.</p> <p>If the romaine was cut by the vendor, collect wash records of the romaine.</p> <p>For all lots identified, please provide grower information and specific field coordinates or farm information, if available.</p> <p>Please provide a customer list of romaine lettuce during the timeframe of interest.</p>
Receiving, ingredient, process	<p>How is romaine lettuce stored, processed, shipped, and handled before delivery? Please provide any relevant safe operating practices (SOPs) detailing practices if available.</p> <p>Does the firm have supplier agreements or contracts down to the grower or farm?</p> <p>Does the firm visit any of the growers or suppliers to observe their growing or processing?</p> <p>Determine how the product of interest is unloaded (is it left outside for any period, who unloads the items; when the items are unloaded, are they immediately placed in their final storage location?)</p> <p>Determine the details of romaine lettuce received at the firm supplied to the POS. Whole head, bagged whole head, prechopped, or other? What type of romaine (variety) is received? What type of romaine was supplied to the retailer?</p> <p>Does raw material arrive in reusable plastic containers? If so, please describe the process for the containers being returned.</p> <p>What size are the containers and how is this reflected on the records?</p> <p>Does the firm conduct any processing or repackaging of the romaine?</p> <p>How does the firm receive, and rotate their stock? What is the turnover rate for romaine lettuce?</p> <p>Is a first-in first-out rotation used?</p> <p>What is the turnaround time from receipt to shipment of romaine lettuce from a processing facility?</p> <p>Are raw lots comingled from different growers into the finished product lots?</p> <p>Is the same equipment used to manufacture different products (list of items manufactured with the same equipment)? How often is this equipment cleaned? What is this equipment cleaned with? Please provide any relevant safe operating practices.</p> <p>Obtain the truck monitoring process for cleanliness of the truck as well as temperature control during transit (and temperature logs).</p>
Additional information requests at the retail level at the farm level	<p>Please refer to the April 2001 traceback procedures, which can be found at <a href="http://www.fda.gov/ICECI/Inspections/InspectionGuides/ucm075005.htm">http://www.fda.gov/ICECI/Inspections/InspectionGuides/ucm075005.htm</a>, for the investigation methodology and reporting requirements. In accordance with this guidance, please also cover the following: the firm's shipping and receiving practices, handling and storage practices, stock rotation practices, daily stock inventory, and order practices, because this information will be needed to determine which shipments were available and most likely consumed at the time the implicated meals were consumed.</p> <p>Fill out form FDA 3623, Farm Investigation Questionnaire. Collect detailed information through interviews, photographs, and supported with documents.</p> <p>Collect a map with ranch, field, or commodity designations. Please mark locations on the map where samples are collected.</p> <p>Collect a water flow map.</p> <p>If possible, please determine where mobile farm equipment that was used during the timeframe of interest is currently or its destination if still en route.</p> <p>If compost is received from another facility, please collect the name of the facility.</p>

tions can provide the additional line of evidence needed to confirm a food source as the source of an outbreak that was suspected based on epidemiologic information. As a traceback investigation progresses, locations for microbiological sampling of foods and production and growing environments are identified, providing additional opportunities to confirm the hypothesis generated by the epidemiologic

investigation. Traceback investigations can also assist in determining the scope of necessary recall actions, public health advisories, import alerts, or other actions to mitigate the risk of an ongoing foodborne illness outbreak. Although traceback can identify the locations where contamination may have happened, further investigation is needed to determine how the contamination happened.

Month	May									
Day	Fri	Sat	Sun	Mon	Tues	Wed	Thu	Fri	Sat	Sun
Date	9	10	11	12	13	14	15	16	17	18
AT: Steak House* City, State										
FROM: XYZ Produce Distributors City, State	8 cs - 6/2lb Chop PO#7894				16 cs - 6/2lb Chop PO#71729			6 cs - 6/2lb Chop PO#71729		12 cs - 6/2lb Chop PO#71723
FROM: Nations Foods Distribution City, State	6CS - 6/2lb Romaine Chop PO #404396		6/2 LB 6CS Romaine Chop PO# 404903							
AT: XYZ Produce Distributors City, State										
FROM: Zenith Fresh Processing City, State				90 cs - 6/2 Romaine Chop PO# 071580		90 cs - 6/2 Romaine Chop PO# 071627		180cs - 6/2 Romaine Chop PO# 071723		
FROM: Superior Veg Processors City, State	80 cs Chop Lot # 734	46 cs Chop Lot # 735								
AT: Zenith Fresh Processing City, State										
FROM: World Farms			Harvest Date: 5/11 Run # 194636 Romaine Grower: ATP Production Date: 5/12	Harvest Date: 5/12 Run # 194635 Romaine Grower: ZPT Production Date: 5/12	Harvest Date: 5/13 Run # 194629 Romaine Grower: MO Production Date: 5/13					

FIGURE 1. An example of a traceback timeline for a multistate foodborne illness outbreak linked to the consumption of romaine lettuce. The source of origin and the destination of each shipment is displayed vertically on the left. The dates of shipments are displayed horizontally on the top. Dates of shipments, product amounts, and lot code information of purchases of romaine lettuce products are traced and displayed in each box from the point of service, through the distribution chain, to processors, distributors, and farms. Implicated shipments that match the timeframe of illness and purchase date (18 May) for the selected timeframe of interest in the boxes are highlighted and arrows display the direction of the shipments from farms, to distributors, to processors, and to the points of service.

Additional environmental investigations, using an array of investigational tools such as environmental assessments (29), are necessary to further determine how contamination may have occurred. An environmental assessment is an investigation to learn what factors may have contributed to an outbreak of foodborne illness or a food contamination event (29). These findings may help identify practices or conditions at the product source that may have led to the contamination event and inform the development of guidelines and policies aimed at minimizing the risk of future contamination and preventing foodborne illnesses. This may also lead to outreach and education efforts to industry and consumers and targeted food surveillance efforts.

Traceback is a major component of a foodborne illness outbreak response, but it is also important to emphasize that like all parts of the outbreak investigation process, it is in response to an event that is already happening or may even already be over. Amid an ongoing foodborne illness outbreak, the overall goal of traceback investigations is to reduce morbidity and mortality by removing contaminated food products from the market. Because of the need for rapid action to prevent additional illnesses, the retrospective nature of traceback investigations and the time needed to conduct them can pose limitations to their utility.

**Traceback interpretation.** Traceback investigations are not a stand-alone tool in identifying the food causing a foodborne illness outbreak. As noted above, a suspected

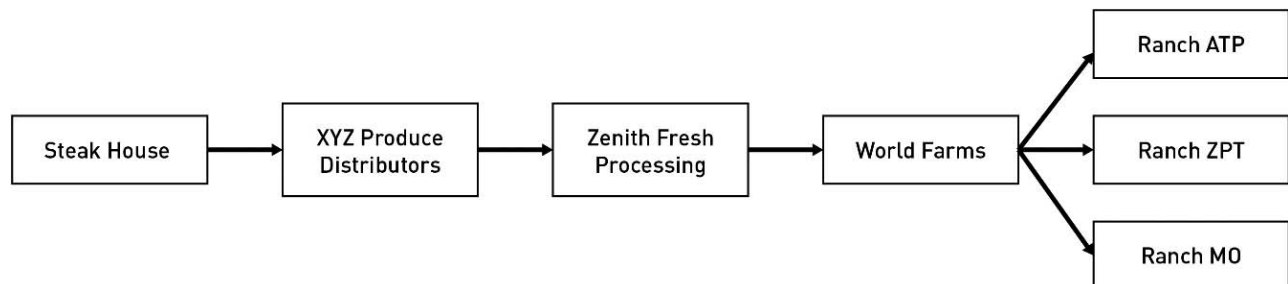


FIGURE 2. An example of a traceback diagram for a multistate foodborne illness outbreak linked to the consumption of romaine lettuce. Purchases of romaine lettuce products are traced from the point of service, through the distribution chain, to processors, distributors, and farms. Product originates from individual ranches, a subdivision of farms, denoted on the right side of the diagram.

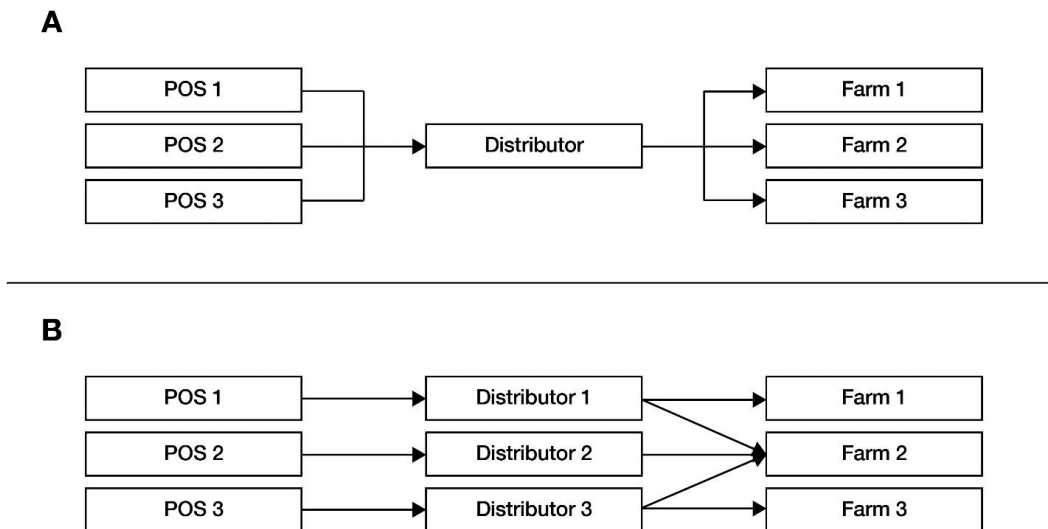


FIGURE 3. Level and strength of convergence identified by traceback investigations in scenario A, where three POSs trace back to a single distributor and three different farms, and scenario B, where three POSs trace back to three distributors and a common farm.

food item identified through the epidemiologic investigation can be confirmed with additional supporting microbiologic or traceback evidence (22). In rare circumstances, the epidemiologic data alone can be compelling enough to confirm a vehicle (e.g., when the brand information reported by ill people in interviews is so specific that it essentially amounts to providing both epidemiologic and traceback evidence of the food source). Foodborne illness outbreak investigations that lack microbiological evidence (e.g., isolating the outbreak strain from a food or environment sample) can still confirm a food as the source of the outbreak by using epidemiologic and traceback investigation data, but both the epidemiologic data and traceback findings need to be conclusive. Generally, any food can be traced through a supply chain at any given time; therefore, it is important that only foods epidemiologically linked to case patients and supplied within a certain timeframe of interest are traced during an outbreak investigation.

Determining the strength of evidence requires extensive analysis of large amounts of supply chain data. One way to determine the strength of the traceback is to assess characteristics about the level of convergence. The level in a supply chain at which convergence occurs has a direct impact on the interpretation of the strength of the finding. This is illustrated by the following example scenarios A and B in Figure 3.

In scenario A, three illness subclusters reported consuming salads containing romaine lettuce at three different POSs. Traceback initiated at each POS identified a common distribution center. Convergence, in this instance, is at the distribution center level. This convergence may be an important lead, but does not necessarily confirm that the distribution center is the source of the romaine contamination, given a distribution center typically moves the romaine between a farm and a POS without ever opening boxes or bags.

The complexity of the supply chain should also be considered. In the aforementioned scenario, each POS is

supplied by the same distribution center, so it is not surprising that they also received romaine lettuce from the same potential farms, but further investigation would be necessary to confirm which farm was the common source.

In scenario B, three POSs were supplied by three different distribution centers, each with its own set of romaine farms in the supply chain. Finding a commonality across these three traceback legs at the farm level improves confidence in the identification of farm 2 as the source. This type of convergence, paired with strong epidemiologic data, can confirm the source of an outbreak, absent microbiological confirmation from a product or environmental sample.

Convergence identified by analysis of supply chain data (such as a single contaminated field or ranch of romaine lettuce or a single processor of mixed greens bagged salad) provides the opportunity to target investigational resources; however, further investigation and sampling may be needed to confirm the convergence as a contributor to the outbreak. Identifying convergence is extremely time-intensive and often hinders the ability of public health professionals to complete traceback investigations before the outbreak is over, particularly for products such as leafy greens, when short shelf life, rapid turnover at the POS, and epidemiologic limitations further hamper timeliness (18, 27). Typically, leafy green harvest and distribution have been completed before the implicated ranch can be identified (24). If an implicated ranch is identified and an on-farm investigation takes place, even in instances when the outbreak strain is identified in the environment, determining the source or route of contamination when the event took place in the past is an ongoing challenge. This is due to several limitations, the most notable of which is that the contamination event(s) took place in the past and the ranch may have moved on to growing or harvesting other crops or is not in production at that time (9, 23).

TABLE 3. Traceback investigation information, including reported illnesses, subcluster selected POS selection, and convergence level for three outbreaks of STEC infections linked to leafy greens from the Yuma growing region, the Santa Maria growing region, and the Salinas Valley growing region

Epidemiologic and traceback information	Outbreak season and year		
	Spring 2018 (Yuma growing region)	Fall 2018 (Santa Maria growing region)	Fall 2019 (Salinas Valley growing region)
Reported illnesses	240	62	172
Hospitalizations	104	23	88
Hemolytic uremic syndrome	28	2	16
Deaths	5	0	0
Total illness subclusters or single case exposures traced	23	6	15
Total farms	12	17	16
Total ranches	26	15	43

### THREE CASE STUDIES OF RECENT FOODBORNE ILLNESS OUTBREAKS OF *E. COLI* O157:H7 INFECTIONS LINKED TO ROMAINE LETTUCE

As described above, the process to initiate, execute, analyze, and take public health actions based on the results of traceback investigations can be complicated and time-intensive, but highly consequential in determining the direction of a foodborne illness outbreak investigation. To illustrate this process, we present an overview of three case studies of recent foodborne illness outbreaks of STEC infections linked to romaine lettuce from the Yuma, Santa Maria, and Salinas growing regions, referred to hereafter as the spring 2018, fall 2018, and fall 2019 outbreaks, respectively. A summary of the epidemiologic and traceback information for the three outbreaks is provided in Table 3. It is important to note that generally, in the United States, from May to November, most romaine lettuce shipments come from California, whereas from December to April, most come from Arizona (1). All three of these foodborne illness outbreaks occurred toward the end of each region's production season. These outbreaks accounted for a total of 474 ill persons that required the mobilization of a vast number of public health professionals at the local, state, and federal level to protect public health. In addition, these three outbreaks traced back to a total of 45 farms and 84 ranches, based on U.S. illness subclusters and single case exposures that were selected for traceback.

**Spring 2018 outbreak.** Starting in April 2018, the FDA, the CDC, and state partners investigated a multistate foodborne illness outbreak of *E. coli* O157:H7 infections. Ultimately, 240 people infected with the outbreak strain were reported from 37 states, 104 people were hospitalized, 28 people developed hemolytic uremic syndrome, and 5 deaths were reported. In response to this foodborne illness outbreak investigation, consumers were warned not to consume romaine lettuce grown in the Yuma growing region (composed of growing areas in Yuma, AZ, and Imperial Valley, CA) while the foodborne illness outbreak was investigated. The traceback investigation was able to identify multiple farms and ranches that converged in the Yuma growing area, leading to the initiation of an

environmental assessment. Water samples collected from an irrigation canal surrounding several ranches identified by the traceback in the region yielded the outbreak strain of STEC O157 (2). Details of the outbreak findings have been reported elsewhere (2, 6, 30, 31).

The traceback investigation served a dual purpose in this foodborne illness outbreak investigation. First, early traceback findings assisted public health officials in generating messaging to inform the public where the contaminated romaine was generally grown. Second, more definitive traceback findings assisted in identification of global positioning system-mapped ranches that led to an environmental assessment, allowing investigators to focus resources on targeted areas. This approach led to successful isolation of the outbreak strain in canal water, a finding that has spurred much research and discussion about future prevention strategies (33).

Initiation of the federal traceback investigation occurred before the illnesses being confirmed as part of an outbreak by CDC PulseNet. The New Jersey Department of Health identified a surge in reports of *E. coli* O157:H7 infections, and quick action by the state health officials led to early identification of a common national restaurant chain as the first illness subcluster (POS A-1 to POS A-8; Fig. 4). Analysis of food exposure data and loyalty card information led to the identification of the POS locations, meal dates, and the suspect ingredient, information that was subsequently shared with the national chain headquarters. Equipped with that information, ingredient analysis of menu items reported by the ill individuals identified romaine lettuce as the common ingredient across the meal items reported. Additional follow-up with the national chain's romaine lettuce processor occurred (processor A; Fig. 4), and preliminary supply chain information was provided to the FDA electronically within approximately 2 days. These electronic records captured the following information at each point in the supply chain: receiver name, supplier name, date of shipment, sell-by date, quantity shipped, product name, sales order number, and customer list. In addition, manufacturing lot code information was provided for outgoing shipments from the processor to the distribution centers; however, lot code



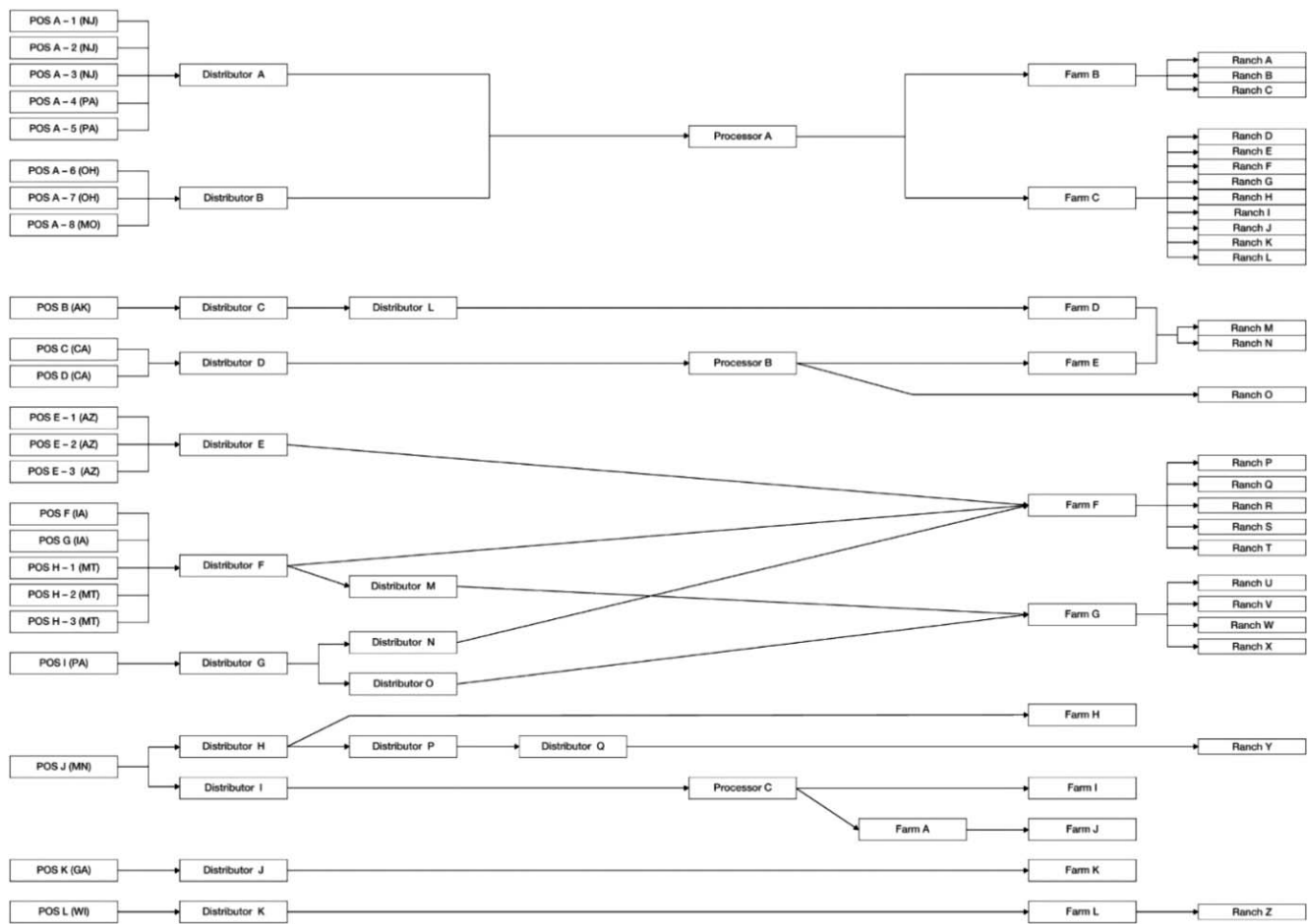


FIGURE 4. Traceback diagram for multistate foodborne illness outbreak of STEC infections linked to the consumption of romaine lettuce from the Yuma growing region in Spring 2018. Purchases of romaine lettuce products are traced from the point of service (POS), through the distribution chain, to processors, distributors, and farms. Product originates from individual ranches, a subdivision of farms, denoted on the right side of the diagram.

information was not further documented from the distribution center to the POS. Although information for the 10 case patients included in the illness subcluster linked to the national chain was collected, additional traceback information from the POS for other illness subclusters was concurrently analyzed. Early receipt of this information in an electronic format provided the ability to determine the Yuma growing region as the source of romaine lettuce to multiple illness subclusters.

After the early identification of the region of interest, the traceback investigation continued to expand. Additional traceback legs were added to the investigation, including a traceback leg identifying an institution (POS B; Fig. 4) that received whole head romaine lettuce supplied by a single farm (farm D; Fig. 4). Although this farm was not identified as the common source across the overall outbreak, identifying a sole supplier for this illness subcluster significantly aided the FDA's ability to focus the ensuing environmental assessment in the Yuma growing region.

The traceback investigation determined that romaine lettuce was supplied to all POSs through 17 distributors and 3 processors (Fig. 4). Except for the whole head romaine lettuce supplied to POS B, all romaine lettuce was comingled. Romaine lettuce from six farms (farms A, B,

C, E, I, and J; Fig. 4) and ranch O was comingled at processors A, B, and C; the remaining six farms comingled romaine lettuce from multiple ranches before distribution. Overall, the traceback investigation identified 9 traceback legs composed of 23 POSs representing 44 case patients. Romaine lettuce consumed at the POS was sourced from 12 farms receiving leafy greens from 26 ranches identified across the Yuma growing region. Comingling and lack of lot coding information prevented further narrowing of the number of farms and ranches identified.

**Fall 2018 outbreak.** Starting in November 2018, the FDA, the CDC, and state partners began investigating a multistate foodborne illness outbreak of *E. coli* O157:H7 infections detected by CDC PulseNet. Ultimately, 62 people infected with the outbreak strain were reported from 16 states and the District of Columbia, 23 people were hospitalized, and 2 people developed hemolytic uremic syndrome; no deaths were reported. In response to this foodborne illness outbreak, consumers were initially warned not to eat any romaine lettuce on the market (21). As the investigation continued, the advice to consumers was ultimately narrowed to only avoid romaine lettuce from three California counties including the Santa Maria growing

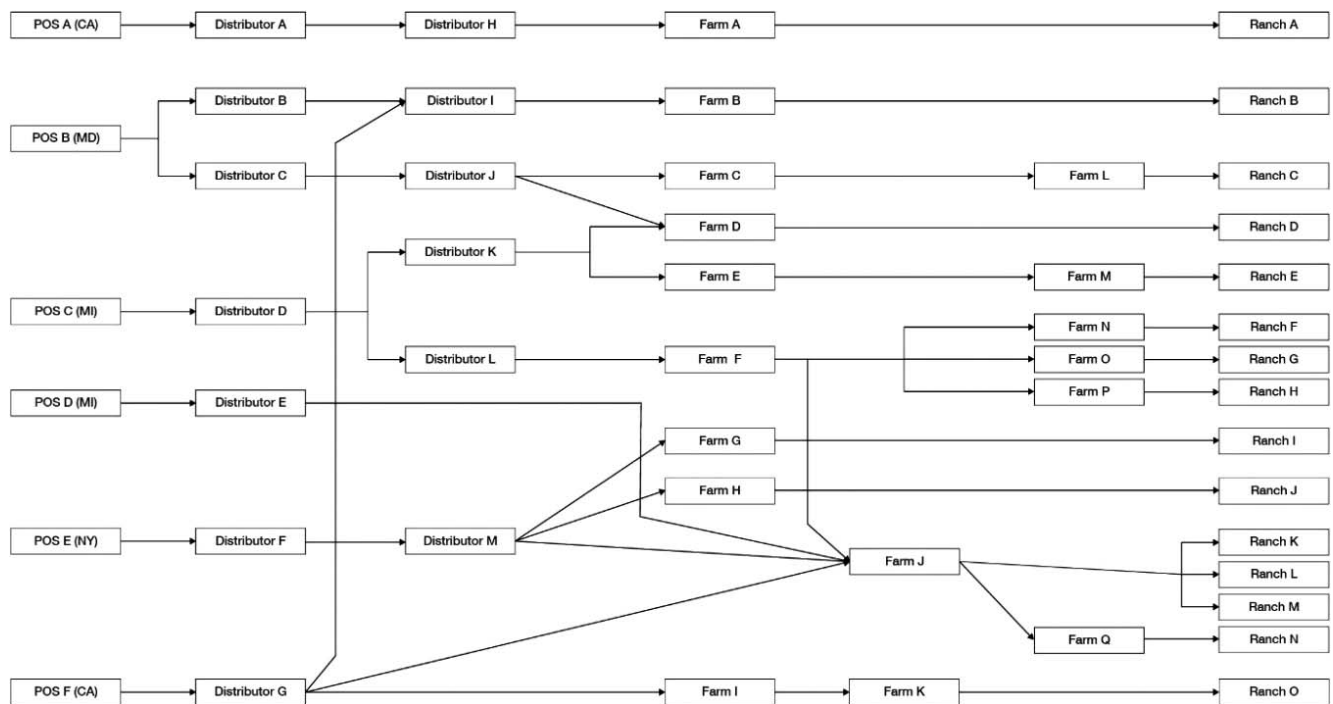


FIGURE 5. Traceback diagram for multistate foodborne illness outbreak of STEC infections linked to the consumption of romaine lettuce from the Santa Maria growing region in fall 2018. Purchases of romaine lettuce products are traced from the point of service (POS), through the distribution chain, to processors, distributors, and farms. Product originates from individual ranches, a subdivision of farms, denoted on the right side of the diagram.

region. An on-farm investigation conducted at one of the farms identified in the traceback revealed the presence of the outbreak strain of *E. coli* O157:H7 in sediment from a reservoir. Details of the outbreak findings have been reported elsewhere (32, 34).

Although romaine was a suspect vehicle initially, the CDC and state health departments could not rule out hummus at the start of the investigation. Traceback of both romaine lettuce and hummus was initiated for illness subclusters reporting salad exposures at local restaurants, as well as purchase information from regional grocery stores. Additional epidemiologic data collected by local and state officials ruled out hummus as a suspect food, and the Coordinated Outbreak Response and Evaluation Network worked with state partners to continue traceback efforts on romaine lettuce.

First, the traceback investigation allowed government officials to narrow the scope of the public health warning from all romaine lettuce on the market to specifically romaine lettuce grown in three California counties that make up the Santa Maria growing region: Monterey, San Benito, and Santa Barbara. Second, the traceback investigation allowed federal and state partners to focus on-farm investigational resources to certain farms and ranches that supplied product to multiple POSs. As noted above, one of these investigations led to the successful isolation of the outbreak strain from an agricultural water reservoir, a finding that led to follow-up investigations (32).

Records received ranged from electronic spreadsheets linking sold product to purchased product and farm information to photocopies of hand-written receipts with

only the immediate supplier name and total amount of romaine purchased and price during the timeframe of interest. Records from four local restaurant illness subclusters (POSs A, B, C, and E; Fig. 5) were hand-written or were paper documents that contained limited data on the source of romaine coming into the facilities during the timeframe of interest. From these paper records, investigators were only able to discern the amount of romaine that was billed to the POS on a specific day by a specific vendor, resulting in a time-intensive and step-wise record collection process through the supply chain to obtain grower information for the potential shipments. All restaurants were supplied by different wholesale produce terminal markets (distributors B, H, J, K, and M; Fig. 5). Records from two grocery store chains (POSs D and F; Fig. 5) were obtained in sortable spreadsheets; however, only one of the distributors (distributor E; Fig. 5) was able to link incoming shipments with lettuce supplied to the store. Without a common identifier of product linking back to a farm or ranch, the FDA, in collaboration with state partners, attempted traceback record collection at more than 30 firms to identify the farms and ranches where the romaine was sourced. As more information was obtained and reviewed, the scope of the public warning was narrowed.

The final FDA traceback investigation, completed with assistance from local and state officials, was composed of six POSs representing eight case patients and identified 13 distributors and 17 farms. In total, 15 ranches were identified in Santa Barbara (ranches A, D, E, I, and O; Fig. 5), Monterey (ranches B, C, G, H, J, K, L, M, and N; Fig. 5), and San Benito (ranch F; Fig. 5), California,

counties within the Santa Maria growing region as potentially supplying romaine lettuce to the POS identified by ill persons during the timeframe of interest. The FDA and state partners conducted sampling and on-farm investigations at multiple ranches identified in the traceback investigation. Only one investigation yielded a sample with the outbreak strain (farm D, ranch D; Fig. 5). The sample collected was sediment of an on-farm water reservoir used by the farm. Identification of the outbreak strain in a single environmental sample does not show how products from the farm became contaminated, as reported elsewhere, in detail (34). During the timeframe of interest, ranch D supplied the leafy greens to two POSs (B and C; Fig. 5). Although no single entity could account for all illnesses in the outbreak, the traceback investigation was successful in identifying a farm where the outbreak strain was found.

**Fall 2019 outbreak.** Starting in November 2019, the FDA, the CDC, and state partners investigated a multistate foodborne illness outbreak of *E. coli* O157:H7 infections. Ultimately, 172 people infected with the outbreak strain were reported from 28 states; 88 people were hospitalized, 16 people developed hemolytic uremic syndrome, and no deaths were reported. In response to this foodborne illness outbreak, consumers were warned not to consume romaine lettuce grown in the Salinas Valley growing region, in California, while the outbreak was investigated. Microbiological data played a key role in this outbreak. By the end of the investigation, the outbreak strain was detected from two separate romaine products and was also found in the environment close to where those leafy greens were grown. Details of the outbreak have been described elsewhere (7, 35, 38). During this foodborne illness outbreak investigation, the FDA, the CDC, and state partners were also responding to two other concurrent foodborne illness outbreaks of *E. coli* O157:H7 linked to leafy greens that were not related to isolates in this outbreak by whole genome sequencing (15, 34, 41). The traceback investigation for the larger multistate outbreak is described below, but it is important to mention the additional outbreaks because they also traced back to the Salinas Valley growing region, and follow-up field investigations were planned with information from all three outbreak investigations.

Before identifying an outbreak vehicle for the multistate investigation, Wicomico County Health Department in Maryland collected, and the Maryland Department of Health tested, an intact prepackaged salad product that was reported by ill persons in their state purchased from various locations of the same chain of grocery stores (POSs A-1, A-2, and A-3; Fig. 6). The romaine lettuce component of the salad yielded the outbreak strain of *E. coli*. This finding prompted a traceback of the romaine lettuce used in that production lot, resulting in the identification of three farms and four ranches (farms A, B, and C; ranches A, B, M, and N; Fig. 6). Later in the investigation, a second intact prepackaged salad that was purchased from POS C (Fig. 6) was tested by the Wisconsin Department of Agriculture's Trade and Consumer Protection, Bureau of Laboratory

Services, and yielded the outbreak strain of *E. coli*. Traceback of this production lot identified two farms and four ranches (farms A and F; ranches C, D, K, and U; Fig. 6). The traceback investigation determined that romaine lettuce in prepackaged salad collected in these two states was sourced from a common farm, but no common ranches.

As information was collected on the Maryland-collected sample that yielded the outbreak strain, illness subclusters and single case exposures across multiple states reporting romaine consumption were identified for traceback. Although having two product samples that yielded the outbreak strain with lot codes were critical findings, these products and the ranches that supplied them could not explain all the illnesses in the outbreak and the case count continued to increase. Additional illness subclusters and single case exposures (POSs B, D, E, F, and G; Fig. 6) were traced with the goal of learning more about the potential source of contamination.

As with the spring 2018 outbreak linked to romaine lettuce from the Yuma growing region, this traceback investigation played two important roles. First, information obtained by state partners early in the traceback investigation, including farm-level data for the romaine contained in the products that yielded the outbreak strain, allowed public health officials to quickly generate public messaging to inform consumers on what foods from which growing region to avoid, to prevent further illnesses. This public messaging was further enhanced by a voluntary, industry-led labeling initiative specifying the growing region where romaine lettuce was sourced displayed on consumer packages (28). Second, the traceback identified farm and ranch locations that were targeted for onsite follow-up. Multiple on-farm investigations occurred during late 2019 and early 2020 that ultimately led to finding the outbreak strain in a cattle fecal-soil composite sample taken less than 2 mi (3.2 km) from a farm identified in the traceback investigation (38). This finding may be useful in learning more about how contamination happened in this event and how future contamination may be prevented when growing produce.

The romaine products that were ultimately sold at the 15 POSs went through 14 distributors (distributors A to N; Fig. 6) and 4 processing facilities (processors A to D; Fig. 6). The traceback investigation identified 13 domestic romaine farms (farms A to M; Fig. 6) and 40 individual ranches (ranches A to AN; Fig. 6) that supplied romaine to the identified POSs during the timeframe of interest. Three Mexican farms and ranches (farms N to P, ranches AO to AQ; Fig. 6) also supplied romaine to one distributor (distributor N; Fig. 6) during the timeframe of interest. Although no single farm was identified that could explain all POSs traced, a farm (farm A; Fig. 6) was identified as supplying romaine lettuce to all POSs except POS-F. This includes the two POSs associated with the products that yielded the outbreak strain (POS A [1 to 3], C; Fig. 6). The romaine supplied by farm A was sourced from 13 different ranches.

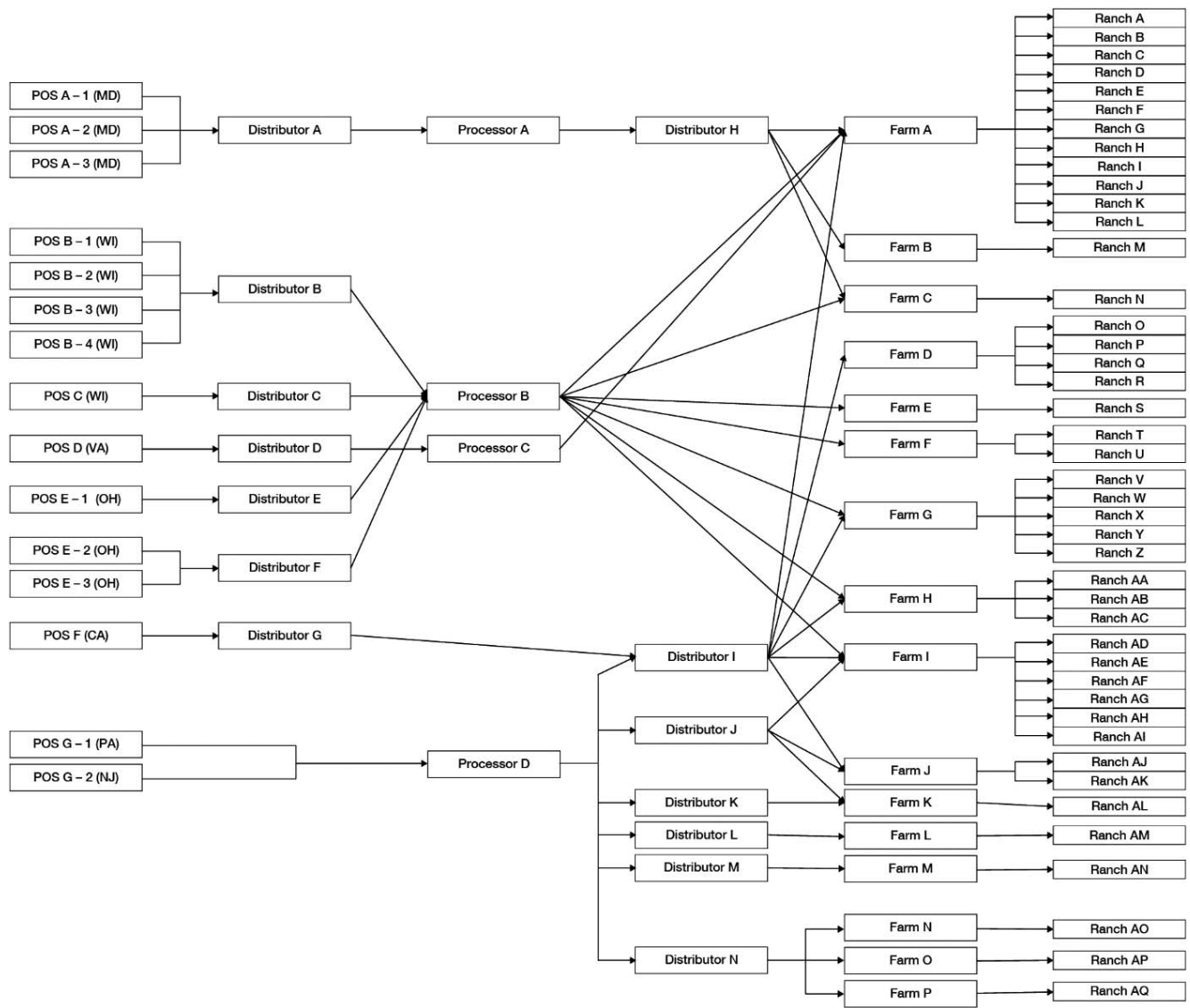


FIGURE 6. Traceback diagram for multistate foodborne illness outbreak of STEC infections linked to the consumption of romaine lettuce from the Salinas growing region in fall 2019. Purchases of romaine lettuce products are traced from the point of service (POS), through the distribution chain, to processors, distributors, and farms. Product originates from individual ranches, a subdivision of farms, denoted on the right side of the diagram.

### CHALLENGES ENCOUNTERED DURING TRACEBACK INVESTIGATIONS LINKED TO LEAFY GREENS

**Initiation of traceback.** One challenge encountered during these traceback investigations has been the availability of epidemiologic information necessary for the initiation of the process. Public health and regulatory bodies rely on information provided by ill people to identify a suspect food source and initiate traceback. Food histories from ill people are obtained through interviews, often weeks or more after initial consumption of contaminated product that may result in poor recall (8, 19). Some consumers impacted by the foodborne illness outbreak may be children, seriously ill, or deceased, and in these situations the food history must be provided by another person, introducing additional uncertainty. Furthermore, not all ill people in outbreaks can be reached for an interview. When

ill people are reached, providing food histories to public health officials is a voluntary process and, in some instances, ill people decline to participate. Relying on an ill person’s recollection of leafy greens in particular, where important details such as types of leafy greens consumed or brand information is often forgotten or unknown, but critical to both the epidemiologic and traceback investigation, can be a significant challenge (18). In addition, multiple instances of leafy greens consumption are often reported by a single patient throughout the exposure period, making it difficult to determine which food exposure caused illness.

The case studies noted above describe traceback investigations linked to the same commodity that were initiated under three different circumstances due to the variability in early exposure data. For the spring 2018 outbreak (Fig. 4), early identification of an illness subcluster



of a chain restaurant (POS A; Fig. 4) with a loyalty card program assisted public health and regulatory officials in noting the specific type of salads consumed by the ill individuals, purchase locations, and dates of purchase, allowing for scoping of additional record requests from the restaurant's supplier. By contrast, for the fall 2018 outbreak, public health and regulatory officials relied on information provided by ill persons about the types of salads and menu items they recalled consuming at predominantly small, independently owned restaurants. For the fall 2019 outbreak, bagged salads purchased at retail chains were the predominate exposure, but limited recollection of specific brands or types of bagged salad and/or multiple exposures or purchases of several brand or types of bagged salad by case patients created challenges in scoping record requests for suppliers. In some instances, the data collected from shopper card records did not match the information reported by case patients. Without specific purchase information, extensive amounts of data were requested covering multiple products across broad timeframes, which resulted in delays in receiving the information of interest and required more time for data analysis, both of which contributed to delays in the traceback investigation. Purchase information from case patient shopper cards, when available, were essential in helping to narrow the list of food of interest and dates purchased. Although shopper cards and other loyalty cards do not provide lot code information, they do assist in verifying purchases and providing important information such as brand and type of product and purchase date, which are often not reported or recalled by ill people.

**Product-identifying information through the supply chain.** For most products, especially leafy greens and other fresh produce, product-identifying information, such as lot codes, product descriptions, and other product identifiers, created at the time of entry into the supply chain is not typically maintained. For example, romaine lettuce harvested from a ranch on a specific date may be assigned a unique lot code, but as that romaine shipment moves through the supply chain and is processed, lot code data either changes or is lost. Facilities, such as distribution centers, tend to assign new product identifiers and lot codes based on their own internal traceability systems without capturing and cross-referencing the information from the source. Most frequently, lot code traceability is lost as product moves between a distribution center and a POS because it is not maintained or recorded at that point in the supply chain. This can be problematic because key information about the source of the product may no longer be available at the POS, further lengthening the time it takes to determine potential sources of contaminated products. Most traceback investigations are initiated from information at the POS; lack of availability of product identifying information requires investigators to take a step-by-step approach to product tracing, rather than being able to trace directly to points in the supply chain where contamination may have occurred.

Lack of product-identifying information at the POS level prevents traceback investigations from focusing on key shipments and instead requires investigators to trace all potential products that could have been available at the POS from the distribution center or other supplier in the timeframe of interest. For example, a POS identified by a traceback investigation may receive romaine lettuce from a single distribution center. Product-identifying information, such as lot codes, may not be maintained on the incoming shipment reports at the POS or the outgoing shipment reports from the distribution center. Because the traceback data available cannot be used to determine which lot codes went to the specific POS, all incoming romaine shipments to the distribution center in the timeframe of interest must be traced back. Identifying shipments of interest based on time captures too many lot codes that were received by the POS and can even capture lot codes that never made it to the POS. For example, if a distribution center received 50 lot codes of romaine lettuce in the timeframe of interest, investigators must now initiate traceback on each of those shipments to try and identify potential sources of the contaminated romaine, because the specific lot codes received by the POS cannot be identified. The inclusion of these additional lot codes dilutes the potential for convergence by identifying farms that may have had no role in the actual product that was available at the POS. Implicating numerous shipments further slows the tracing process, requiring additional time for unnecessary record collection from firms along the supply chain that may not have supplied contaminated product. This ultimately impacts the timing to perform on-farm investigations and effectively scope product action. As a result, product action is broadened to ensure protection of public health.

In each foodborne illness outbreak noted above, lot code data were not available at the POS, with the sole exception being the fall 2019 outbreak where lot code data were identified on the package of an intact product obtained from an ill person that yielded the outbreak strain. For products where lot codes were available, the entire supply chain was identified within 24 h or less compared with POS data where no lot code information was available, and identification of the entire supply chain took between 5 and 25 days. In the spring 2018 outbreak, some supply chain information from the restaurant chain was quickly available and a processor was immediately identified, but because no production lot codes were maintained at POS A (Fig. 4), purchase orders from POS A were provided to processor A to help implicate production lots and identify specific romaine suppliers. For this single traceback leg (Fig. 4), there were approximately 12 production lots (1.8 million lb [816,466 kg]) available from processor A in the timeframe of interest that were traced back, resulting in the identification of two farms. In the fall 2018 outbreak, lack of lot code information available at each POS delayed traceback, extending the need for a broadly scoped consumer warning.

**Record-keeping issues.** Supply chain information is maintained in a wide variety of ways across industry. As

most establishments use inventory management and procurement systems to fulfill present record-keeping requirements, records produced by these systems, whether paper or electronic, must be analyzed and interpreted along each step of the supply chain. Paper records are typically sent by electronic platforms in portable document format or as scanned copies. However, depending on the quantity of records collected, in some situations, physical copies of the paper records are mailed for review; creating significant delays in the traceback investigation. Firms with more advanced tracing systems can provide their data in electronic format, such as sortable spreadsheets.

Dates, product descriptions, product codes, lot codes, and other fundamental data elements important to traceback are not universally defined, further complicating how the information is passed and interpreted along the supply chain. For example, in the fall 2018 outbreak, some addresses on bills of lading from suppliers were P.O. boxes or residences for the shipper, broker, or delivery service, not the location that the actual product was shipped from or grown. In at least one traceback leg, this led to a delay of 13 days to locate the owner and obtain information on the source of the romaine, which further delayed decisions on advisories that warned consumers about which products to avoid.

Outside of record interpretation issues, some firms continue to maintain poor records. The FDA's existing food traceability record-keeping regulation (set forth in 21 CFR Part 1, Subpart J), adopted in 2004 in accordance with section 414 of the Federal Food, Drug, and Cosmetic Act (added as part of the 2002 Bioterrorism Act), requires firms to know and record the immediate previous sources of their food products and ingredients and the immediate subsequent recipients of the products they make and/or distribute ("one step back and one step forward") (26). Although this regulation is incredibly important for traceback investigations, some POSs, such as restaurants and retail locations, are exempt from the requirement. This gap in legal requirement has made it particularly difficult for foodborne illness outbreak investigations because many tracebacks rely on initiation from these exempted POSs. In many cases, limited information or poor record keeping has prevented or delayed initiation of the investigation. For example, in the fall 2018 outbreak, a distributor (distributor L; Fig. 5) provided records that included hand-written receipts of product sold to a customer and an official memo stating the supplier was a farm (farm F; Fig. 5) with a P.O. box located in an area other than the harvesting farm.

Poor records can lead to dead ends when they are of poor quality and difficult to decipher, or they can represent false information due to human error and lead the investigation in the wrong direction (3). For example, during the fall 2018 outbreak response efforts, growers had maintained records to link sales to harvest locations and dates with lot and brand information for whole head romaine lettuce. However, after being sold through two or three distributors, some POSs received hand-written receipts identifying only the type of lettuce, quantity, and price. As a result, the information collected from the one step back and

one step forward record-keeping requirement for the supply chain was inadequate in this instance, created delays in identifying a growing region for the product, and subsequently led to the sweeping consumer recommendations on which products to avoid.

**Comingling.** For commodities such as romaine and other leafy greens, certain industry practices create challenges for traceback investigations. Comingling of product from multiple suppliers into a repackaged product is a key challenge faced in leafy green tracebacks. A single production lot of bagged romaine likely contains romaine lettuce sourced from multiple farms and/or fields. This complicates traceback investigations because the need to traceback several production lot codes coupled with numerous farms and fields supplying each lot can result in the identification of many endpoints, making it difficult to determine a contamination source. During the fall 2019 outbreak response efforts, a sample that yielded the outbreak strain collected from a bagged salad mix identified by Maryland state partners allowed government officials to provide a specific production code to industry for tracing. Although this led to the rapid identification of the supply chain for the product, the result was the implication of four different ranches from multiple farms, rather than just one ranch. Similarly, the spring 2018 outbreak traceback investigation led to a processor who used romaine lettuce sourced from 2 farms and 12 ranches during the timeframe of interest, preventing linkage of a single source to a single production lot.

Using multiple suppliers in a single production lot increases the number of endpoints identified in the traceback investigation and the scope of the impact of contaminated product. The endpoints of traceback investigations in outbreaks caused by contaminated romaine lettuce are the fields where product was grown and the likely source of STEC contamination. When multiple fields are identified, determining which field to allocate on-farm investigational resources becomes challenging. The more fields that are identified by the traceback, the less likely investigators can narrow which fields contributed to the contamination and thereby prevent the completion of an environmental assessment. When the root cause of these foodborne illness outbreaks is not understood, little can be done to prevent recurrences.

## OPPORTUNITIES FOR IMPROVEMENT AND THE FUTURE OF TRACEBACK INVESTIGATIONS

In the past few years, numerous outbreaks of STEC infections linked to leafy greens have resulted in expansive advisories to not eat, serve, or sell either all romaine lettuce or all romaine lettuce grown in a region (30, 32). One reason for these broad advisories is a lack of uniform data collection across the supply chain. Although the traceback activities successfully identified farms that could have supplied affected product during the timeframe of interest for the case studies, more can and should be done to improve the process to make these investigations more rapid and precise. More rapid traceability systems can help

prevent illnesses by more quickly identifying the distribution of the contaminated product and aiding in more narrow public health advisories and recalls. More precise systems can enhance the ability to identify which farms actually supplied the contaminated product and increase the likelihood of root cause identification. Three key initiatives underway at the FDA to address challenges with traceback investigations are the 2020 Leafy Greens STEC Action Plan, development of new regulations under the Food Safety Modernization Act Section 204, and the New Era of Smarter Food Safety (36, 40).

**2020 Leafy Greens STEC Action Plan.** Repeated foodborne illness outbreaks of STEC infections linked to leafy greens has led the FDA to develop the 2020 Leafy Greens STEC Action Plan (36), which seeks to advance work in prevention, response, and knowledge gaps concerning STEC contamination of leafy greens. Specific to challenges identified during recent outbreak investigations, the plan calls for the FDA to work with retailers and government partners to improve how shopper card data are requested and shared during outbreak response efforts. The creation of data-sharing templates to standardize information shared between retailers and government could streamline the fulfillment of traceback data needed from consumer purchases. The plan also calls for the promotion of technology-enabled traceability, spearheaded by the development of pilot studies to test the interoperability of tracing systems across the supply chain, such as a recently published pilot study on leafy green traceability (11).

**Food Safety Modernization Act Section 204(d)(1)—Proposed Rule for Food Traceability.** Food Safety Modernization Act 204(d)(1) seeks to establish the framework of information required to trace specific foods across the U.S. food supply chain. This would be accomplished by establishing a consistent approach for product tracing and defining the fundamental key data elements that firms must establish, maintain, and share with downstream customers in their supply chain. Establishment of the regulation will help provide a basis for consistent traceback terminology, an opportunity for electronic record keeping, and a universal understanding of required key data elements. These improvements to existing record-keeping regulations will assist in building the foundation needed for rapid, interoperable tracing systems for specific commodities.

**New Era for Smarter Food Safety Blueprint.** The FDA has developed a strategic blueprint that outlines how the FDA will leverage technology, and other tools, to create a more digitally traceable and safer food system (40). One of the core elements identified in this blueprint is the need for technology-enabled traceability. This includes the need to develop the foundational components of tracing by standardizing critical tracking events and key data, allowing for interoperability and harmonization across the system. In addition, this element recognizes the need to encourage and incentivize industry adoption of new technologies. Finally,

this challenges the FDA to leverage the digital transformation to improve the use of digital traceability system in outbreak investigation and recall protocols. The New Era for Smarter Food Safety lays out a vision for the future of food traceability. Food Safety Modernization Act Section 204(d)(1) is a key tool in laying the foundation for terminology and interoperability. Upon implementation of tracing technologies across the industry, an immediate impact on outbreak investigations, recall events, and food security issues is envisioned. Future use of digitized supply chain could include predictive modeling and forecasting for food safety and supply chain issues.

## CONCLUSIONS

Traceback investigations provide information critical to solving foodborne illness outbreak investigations. As outlined in case studies of three outbreaks of STEC infections linked to romaine lettuce in 2018 and 2019, traceback investigations provided information to target appropriate resources to identify potential root causes, take public health actions, and inform important initiatives such as the 2020 Leafy Greens STEC Action Plan. Although the success of these traceback investigations highlight the importance of this investigational tool, there is much room for improvement. Specific to leafy greens, there were challenges with case patient food recall information, lack of product-identifying information available at the POS, and product comingling across the outbreaks. Federal and state investigators have developed strategies, such as using shopper card information and specifying timeframe of interest in the absence of product identifying information, to aid in overcoming information gaps. However, implementation of these strategies still takes a significant amount of time to accomplish. More work is needed to improve and standardize the use of these strategies so they can be rapidly conducted, improve the speed at which sources are identified, and achieve the goal of preventing additional illnesses. Electronic supply chain traceability systems currently in use or under development by certain sectors of the food industry are changing the landscape for management of supply chains. Engagement between public health and industry on data captured in these systems could transform the way traceback investigations are performed in the future. Until that time, investigators are currently in a period of transitional tracing, where they need to begin to prepare to digest the digital data available in some supply chains, yet still conduct traceback using traditional paper systems. Future work via action plans, regulations, and strategic planning are essential to creating solutions that address the challenges highlighted by leafy green traceback investigations and move to a more rapid, streamlined, and functional traceback investigation process that operates across government sectors and industries.

## ACKNOWLEDGMENTS

The tireless efforts and assistance from state partners in outbreak response and traceback investigations are highly appreciated, most notably our public health and regulatory partners from Arizona, California, Maryland, Michigan, New Jersey, New York, and Wisconsin. In particular, we thank the Wisconsin Rapid Response Team, the Wisconsin Department



of Agriculture, Trade and Consumer Protection, Bureau of Laboratory Services; the Wisconsin State Laboratory of Hygiene; the La Crosse County Department of Health; the Maryland Rapid Response Team, Maryland Department of Health; and the Wicomico County Department of Health. Special thanks to the FDA's emergency response coordinators and laboratories for their assistance with outbreak response and coordination efforts and to Katie Vierk, Andrew Kennedy, Kurt Nolte, and Brittany Nork for their review and feedback.

## REFERENCES

- Astill, G. M., F. Kuchler, J. E. Todd, and E. T. Page. 2020. Shiga toxin-producing *Escherichia coli* (STEC) O157:H7 and romaine lettuce: source labeling, prevention, and business. *Am. J. Public Health* 110:322–328.
- Bottichio, L., A. Keaton, D. Thomas, T. Fulton, A. Tiffany, A. Frick, M. Mattioli, A. Kahler, J. Murphy, M. Otto, A. Tesfai, A. Fields, K. Kline, J. Fiddner, J. Higa, A. Barnes, F. Arroyo, A. Salvatierra, A. Holland, W. Taylor, J. Nash, B. M. Morawski, S. Correll, R. Hinnenkamp, J. Havens, K. Patel, M. N. Schroeder, L. Gladney, H. Martin, L. Whitlock, N. Dowell, C. Newhart, L. F. Watkins, V. Hill, S. Lance, S. Harris, M. Wise, I. Williams, C. Basler, and L. Gieraltowski. 2020. Shiga toxin-producing *E. coli* infections associated with romaine lettuce—United States, 2018. *Clin. Infect. Dis.* 71:e323–e330.
- Buss, B. F., M. V. Joshi, A. L. O'Keefe, C. D. Allensworth, A. Garvey, K. Obbink, S. Mandernach, and T. J. Safrank. 2016. Regional investigation of a cyclosporiasis outbreak linked to imported romaine lettuce – Nebraska and Iowa, June–August 2013. *Epidemiol. Infect.* 144:1807–1817.
- Carstens, C. K., J. K. Salazar, and C. Darkoh. 2019. Multistate outbreaks of foodborne illness in the United States associated with fresh produce from 2010 to 2017. *Front. Microbiol.* 10:2667.
- Centers for Disease Control and Prevention. 2016. PulseNet. Available at: <https://www.cdc.gov/pulsenet/index.html>. Accessed 11 November 2020.
- Centers for Disease Control and Prevention. 2018. Multistate outbreak of *E. coli* O157:H7 infections linked to romaine lettuce. Available at: <https://www.cdc.gov/ecoli/2018/o157h7-04-18/index.html>. Accessed 6 May 2020.
- Centers for Disease Control and Prevention. 2020. Outbreak of *E. coli* infections linked to romaine lettuce. Available at: <https://www.cdc.gov/ecoli/2019/o157h7-11-19/index.html>. Accessed 21 May 2020.
- Council to Improve Foodborne Outbreak Response. 2020. Guidelines for foodborne disease outbreak response. Available at: <https://cifor.us/downloads/clearinghouse/CIFOR-Guidelines-Complete-third-Ed-FINAL.pdf>. Accessed 29 July 2020.
- Firestone, M., K. Hoelzer, C. Conroy, C. Hedberg, and J. Guzewich. 2018. Leveraging current opportunities to communicate lessons learned from root cause analysis to prevent foodborne illness outbreaks. *Food Prot. Trends* 38:134–138.
- Heiman, K. E., R. K. Mody, S. D. Johnson, P. M. Griffin, and L. H. Gould. 2015. *Escherichia coli* O157 outbreaks in the United States, 2003–2012. *Emerg. Infect. Dis.* 21:1293–1301.
- Institute of Food Technologists. 2020. Leafy greens action plan tech enabled traceability 2020 leafy green pilot final report. Available at: <https://www.ift.org/-/media/gftc/pdfs/fda-leafy-green-pilot-final-report-12220.pdf>. Accessed 17 December 2020.
- Interagency Food Safety Analytics Collaboration. 2019. Foodborne illness source attribution estimates for 2017 for *Salmonella*, *Escherichia coli* O157, *Listeria monocytogenes*, and *Campylobacter* using multi-year outbreak surveillance data, United States. Available at: <https://www.cdc.gov/foodsafety/ifsac/pdf/P19-2017-report-TriAgency-508-revised.pdf>. Accessed 29 July 2020.
- Interagency Food Safety Analytics Collaboration. 2020. Interagency food safety analytics collaboration. Foodborne illness source attribution estimates for 2018 for *Salmonella*, *Escherichia coli* O157, *Listeria monocytogenes*, and *Campylobacter* using multi-year outbreak surveillance data. Available at: <https://www.cdc.gov/foodsafety/ifsac/annual-reports.html>. Accessed 8 March 2021.
- Joseph, A., A. Cointe, P. Mariani Kurkdjian, C. Rafat, and A. Hertig. 2020. Shiga toxin-associated hemolytic uremic syndrome: a narrative review. *Toxins (Basel)* 12:67.
- King County Department of Public Health. 2020. Shiga toxin-producing *E. coli* O157:H7 (STEC) outbreak associated with 7 Evergreens restaurants. Available at: <https://www.kingcounty.gov/depts/health/communicable-diseases/disease-control/outbreak/evergreens.aspx>. Accessed 11 August 2020.
- Kintz, E., L. Byrne, C. Jenkins, C. N. Mc, R. Vivancos, and P. Hunter. 2019. Outbreaks of shiga toxin-producing *Escherichia coli* linked to sprouted seeds, salad, and leafy greens: a systematic review. *J. Food Prot.* 82:1950–1958.
- Laughlin, M., L. Bottichio, J. Weiss, J. Higa, E. McDonald, R. Sowadsky, D. Fejes, A. Saupe, G. Provo, S. Seelman, J. Concepción-Acevedo, and L. Gieraltowski. 2019. Multistate outbreak of *Salmonella* Poona infections associated with imported cucumbers, 2015–2016. *Epidemiol. Infect.* 147:e270.
- Marshall, K., A. Hexemer, S. Seelman, M. Fatica, T. Blessington, M. Hajmeer, H. Kisselburgh, R. Atkinson, K. Hill, D. Sharma, M. Needham, V. Peralta, J. Higa, K. Blickenstaff, I. Williams, M. Jung, M. Wise, and L. Gieraltowski. 2020. Lessons learned from a decade of investigations of shiga toxin-producing *Escherichia coli* outbreaks linked to leafy greens, United States and Canada. *Emerg. Infect. Dis.* 26:2319.
- Mba-Jonas, A., W. Culpepper, T. Hill, V. Cantu, J. Loera, J. Borders, L. Saathoff-Huber, J. Nsubuga, I. Zambrana, S. Dalton, I. Williams, and K. P. Neil. 2018. A multistate outbreak of human *Salmonella* Agona infections associated with consumption of fresh, whole papayas imported from Mexico–United States, 2011. *Clin. Infect. Dis.* 66:1756–1761.
- Miller, B. D., C. E. Rigdon, J. Ball, J. M. Rounds, R. F. Klos, B. M. Brennan, K. D. Arends, P. Kennelly, C. Hedberg, and K. E. Smith. 2012. Use of traceback methods to confirm the source of a multistate *Escherichia coli* O157:H7 outbreak due to in-shell hazelnuts. *J. Food Prot.* 75:320–327.
- Minor, T., G. Astill, S. Raszap, S. Thornsby, J. C. Buzby, C. Hitaj, L. Kantor, F. Kuchler, B. Ellison, and A. K. Mishra. 2020. Economic drivers of food loss at the farm and pre-retail sectors: a look at the produce supply chain in the United States. Economic Information Bulletin No. 216. U.S. Department of Agriculture, Economic Research Service.
- Pightling, A. W., J. B. Pettengill, Y. Luo, J. D. Baugher, H. Rand, and E. Strain. 2018. Interpreting whole-genome sequence analyses of foodborne bacteria for regulatory applications and outbreak investigations. *Front. Microbiol.* 9:1482.
- Selman, C. A., and J. J. Guzewich. 2014. Public health measures: environmental assessment in outbreak investigations, p. 98–106. In Y. Motarjemi (ed.), *Encyclopedia of food safety*, vol. 4. Academic Press, Waltham, MA.
- Slayton, R. B., G. Turabelidze, S. D. Bennett, C. A. Schwensohn, A. Q. Yaffee, F. Khan, C. Butler, E. Trees, T. L. Ayers, M. L. Davis, A. S. Laufer, S. Gladbach, I. Williams, and L. B. Gieraltowski. 2013. Outbreak of Shiga toxin-producing *Escherichia coli* (STEC) O157:H7 associated with romaine lettuce consumption, 2011. *PLoS ONE* 8:e55300.
- Sodha, S. V., M. Lynch, K. Wannemuehler, M. Leeper, M. Malavet, J. Schaffzin, T. Chen, A. Langer, M. Glenshaw, D. Hoefler, N. Dumas, L. Lind, M. Iwamoto, T. Ayers, T. Nguyen, M. Biggerstaff, C. Olson, A. Sheth, and C. Braden. 2011. Multistate outbreak of *Escherichia coli* O157:H7 infections associated with a national fast-food chain, 2006: a study incorporating epidemiological and food source traceback results. *Epidemiol. Infect.* 139:309–316.
- Taylor, E., J. Kastner, and D. Renter. 2010. Challenges involved in the *Salmonella* Saintpaul outbreak and lessons learned. *J. Public Health Manag. Pract.* 16:221–231.
- Turner, K., C. N. Moua, M. Hajmeer, A. Barnes, and M. Needham. 2019. Overview of leafy greens-related food safety incidents with a California link: 1996 to 2016. *J. Food Prot.* 82:405–414.
- United Fresh Produce Association. 2019. Questions & answers on voluntary romaine growing region labeling May 7, 2019. Available



- at: <https://www.unitedfresh.org/content/uploads/2019/05/Question-and-Answers-on-Romaine-Outbreak-Winter-2019-7May-002.pdf>. Accessed 29 January 2021.
29. U.S. Food and Drug Administration. 2018. Environmental assessments from foodborne illness or contamination events. Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/environmental-assessments-foodborne-illness-or-contamination-events>. Accessed 6 May 2020.
  30. U.S. Food and Drug Administration. 2018. FDA investigated multistate outbreak of *E. coli* O157:H7 infections linked to romaine lettuce from Yuma growing region. Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/fda-investigated-multistate-outbreak-e-coli-o157h7-infections-linked-romaine-lettuce-yuma-growing>. Accessed 6 May 2020.
  31. U.S. Food and Drug Administration. 2018. FDA update on traceback related to the *E. coli* O157:H7 outbreak linked to romaine lettuce. Available at: <https://www.fda.gov/news-events/fda-voices/fda-update-traceback-related-e-coli-o157h7-outbreak-linked-romaine-lettuce>. Accessed 21 May 2020.
  32. U.S. Food and Drug Administration. 2018. Outbreak Investigation of *E. coli*: romaine (November 2018). Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/outbreak-investigation-e-coli-romaine-november-2018>. Accessed 6 May 2020.
  33. U.S. Food and Drug Administration. 2019. FDA partners with the University of Arizona, Wellton-Mohawk irrigation and drainage district, and Yuma area leafy greens stakeholders to enhance food safety. Available at: <https://www.fda.gov/food/cfsan-constituent-updates/fda-partners-university-arizona-wellton-mohawk-irrigation-and-drainage-district-and-yuma-area-leafy>. Accessed 12 May 2020.
  34. U.S. Food and Drug Administration. 2019. Investigation summary: factors potentially contributing to the contamination of romaine lettuce implicated in the fall 2018 multi-state outbreak of *E. coli* O157:H7. Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/investigation-summary-factors-potentially-contributing-contamination-romaine-lettuce-implicated-fall>. Accessed 12 May 2020.
  35. U.S. Food and Drug Administration. 2019. Outbreak Investigation of *E. coli*: romaine from Salinas, California (November 2019). Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/outbreak-investigation-e-coli-romaine-salinas-california-november-2019>. Accessed 6 May 2020.
  36. U.S. Food and Drug Administration. 2020. 2020 Leafy greens STEC action plan. Available at: <https://www.fda.gov/food/foodborne-pathogens/2020-leafy-greens-stec-action-plan>. Accessed 14 August 2020.
  37. U.S. Food and Drug Administration. 2020. About the CORE Network. Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/about-core-network>. Accessed 5 May 2020.
  38. U.S. Food and Drug Administration. 2020. Factors potentially contributing to the contamination of romaine lettuce implicated in the three outbreaks of *E. coli* O157:H7 during the fall of 2019. Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/factors-potentially-contributing-contamination-romaine-lettuce-implicated-three-outbreaks-e-coli>. Accessed 21 May 2020.
  39. U.S. Food and Drug Administration. 2020. FSMA final rule on produce safety standards for the growing, harvesting, packing, and holding of produce for human consumption. Available at: <https://www.fda.gov/food/food-safety-modernization-act-fsma/fsma-final-rule-produce-safety>. Accessed 6 May 2020.
  40. U.S. Food and Drug Administration. 2020. New era of smarter food safety. Available at: <https://www.fda.gov/food/new-era-smarter-food-safety>. Accessed 14 August 2020.
  41. U.S. Food and Drug Administration. 2020. Outbreak investigation of *E. coli*: salad mix (December 2019). Available at: <https://www.fda.gov/food/outbreaks-foodborne-illness/outbreak-investigation-e-coli-salad-mix-december-2019>. Accessed 11 August 2020.