Shiga Toxin–Producing *E. coli* Infections Associated with Flour

Samuel J. Crowe, Ph.D., M.P.H., Lyndsay Bottichio, M.P.H.,
Lauren N. Shade, B.S., Brooke M. Whitney, Ph.D., Nereida Corral, M.P.H.,
Beth Melius, M.N., M.P.H., Katherine D. Arends, M.P.H.,
Danielle Donovan, M.S., Jolianne Stone, M.P.H., Krisandra Allen, M.P.H.,
Jessica Rosner, M.P.H., Jennifer Beal, M.P.H., Laura Whitlock, M.P.H.,
Anna Blackstock, Ph.D., June Wetherington, M.S., Lisa A. Newberry, Ph.D.,
Morgan N. Schroeder, M.P.H., Darlene Wagner, Ph.D., Eija Trees, D.V.M., Ph.D.,
Stelios Viazis, Ph.D., Matthew E. Wise, M.P.H., Ph.D.,
and Karen P. Neil, M.D., M.S.P.H.

**ABSTRACT**

**BACKGROUND**

In 2016, a multijurisdictional team investigated an outbreak of Shiga toxin–producing *Escherichia coli* (STEC) serogroup O121 and O26 infections linked to contaminated flour from a large domestic producer.

**METHODS**

A case was defined as infection with an outbreak strain in which illness onset was between December 21, 2015, and September 5, 2016. To identify exposures associated with the outbreak, outbreak cases were compared with non-STEC enteric illness cases, matched according to age group, sex, and state of residence. Products suspected to be related to the outbreak were collected for STEC testing, and a common point of contamination was sought. Whole-genome sequencing was performed on isolates from clinical and food samples.

**RESULTS**

A total of 56 cases were identified in 24 states. Univariable exact conditional logistic-regression models of 22 matched sets showed that infection was significantly associated with the use of one brand of flour (odds ratio, 21.04; 95% confidence interval [CI], 4.69 to 94.37) and with tasting unbaked homemade dough or batter (odds ratio, 36.02; 95% CI, 4.63 to 280.17). Laboratory testing isolated the outbreak strains from flour samples, and whole-genome sequencing revealed that the isolates from clinical and food samples were closely related to one another genetically. Trace-back investigation identified a common flour-production facility.

**CONCLUSIONS**

This investigation implicated raw flour as the source of an outbreak of STEC infections. Although it is a low-moisture food, raw flour can be a vehicle for foodborne pathogens.
Flour has been a suspected outbreak vehicle for Shiga toxin–producing Escherichia coli (STEC) infections since 2009, when a multistate outbreak of foodborne disease was linked to prepackaged cookie dough.1,2 However, flour was not definitively identified as the source of infection in that outbreak or in subsequent outbreaks of STEC infection. Flour is a raw, minimally processed product intended to be mixed with other ingredients and cooked before consumption. It is a low-water-content ingredient and typically does not support bacterial growth. Nevertheless, pathogenic microorganisms on the wheat or other ingredients in flour can survive the drying process and remain viable in flour for months in a desiccated state.3,4 STEC, which is estimated to cause 265,000 infections in the United States each year, has been identified as one of a group of pathogens that can contaminate flour.5,6 Symptoms typically appear 3 to 4 days after infection and include mild fever, abdominal pain, vomiting, and diarrhea, which is often bloody. The hemolytic–uremic syndrome, a form of kidney failure, also develops in some patients with STEC infection.7

In 2016, a multistate outbreak investigation in the United States linked infection with STEC serogroups O121 and O26 to contaminated flour from a large domestic producer. We describe the epidemiologic, laboratory, and trace-back aspects of the investigation and discuss the public health implications of our findings.

Methods

Overview of the Investigation

In February 2016, the Centers for Disease Control and Prevention (CDC) and state health departments began investigating a cluster of patients who were infected with STEC O121. All the patients were infected with a strain of STEC O121 that had the same uncommon pulsed-field gel electrophoresis (PFGE) pattern combination, which suggested a common source of illness. CDC and state and local health officials interviewed the patients to obtain demographic, clinical, and exposure information. As the investigation developed, patients who were infected with STEC O121 characterized by other PFGE pattern combinations, as well as patients infected with STEC that had a PFGE pattern of an additional serogroup, O26, were identified as part of the outbreak. After the investigation, members of the investigation team prepared this report for submission; the authors vouch for the accuracy and completeness of the data collected and of the subsequent analyses.

Case Identification

Cases were identified by PulseNet, the national molecular subtyping network for foodborne disease surveillance.8,9 CDC and state laboratory personnel further subtyped selected clinical isolates by means of whole-genome sequencing. The QIAGEN DNeasy Blood and Tissue Kit (Qia-gen) was used to extract genomic DNA. The DNA libraries were created with the Nextera XT DNA Library Preparation Kit (Illumina). DNA sequencing was performed on the Illumina MiSeq Sequencing System. The Lyve-SET pipeline was used to perform high-quality single-nucleotide polymorphism (SNP) analysis.10 Two criteria were used by CDC epidemiologists to define a case. First, a case was an infection with E. coli serogroup O121 that had PFGE pattern combination EXKX01.0001/EXKA26.0001, EXKX01.0001/EXKA26.0313, EXKX01.0389/EXKA26.0001, or EXKX01.0395/EXKA26.0001 or an infection with E. coli serogroup O26 that had pattern combination EVCX01.2685/EVCA26.1686, with illness onset during the period from December 21, 2015, through September 5, 2016. Second, when data were available, the infecting E. coli strain in a case was found by whole-genome sequencing to be closely related genetically to other isolates from clinical samples or to isolates from flour samples collected during the outbreak investigation.

Hypothesis Generation

On the basis of preliminary interview data collected by local and state officials, CDC epidemiologists created and deployed an initial questionnaire to identify common exposures. The frequencies of exposures were compared with those in the Foodborne Disease Active Surveillance Network (FoodNet) Population Survey — a survey in which data on food-consumption frequencies during the previous 7 days are collected from interviewees — with the use of a binomial probability distribution.11 After the responses to the initial questionnaire did not lead to the generation of a strong hypothesis regarding the outbreak source, a CDC epidemiologist conducted open-ended interviews with a subset of patients. Inquiries were made about all foods consumed during the week before illness onset.
and locations where the foods were eaten. Findings from these interviews led to the development of a second questionnaire that local and state health officials then used to interview additional case patients.

**CASE—CASE ANALYSIS**

CDC and state epidemiologists conducted a matched case–case analysis of the data collected with the use of the second questionnaire to identify exposures associated with illness. Cases from this outbreak were matched with cases of infection caused by non-STEC enteric pathogens, such as salmonella, that had been reported to state health departments during the outbreak period (non-outbreak cases). Outbreak case patients and patients with non-STEC illness were matched according to state of residence, sex, and age group (<1 to 9 years, 10 to 19 years, 20 to 29 years, and ≥30 years). Four cases of non-STEC illness were sought for each outbreak case. Odds ratios and 95% confidence intervals were calculated for food exposures with the use of univariable and multivariable exact conditional logistic-regression models (clogit function from the Survival Analysis package in R software). The most probable outbreak sources on the basis of the results of the univariable analyses were included in the multivariable model.

**PRODUCT TRACE-BACK INVESTIGATION AND TESTING**

Local and state health officials collected data on production lots and “better if used by” dates for foods suspected to be involved in the outbreak to determine whether they could be traced to a common production location and time period. The Food and Drug Administration (FDA) inspected production locations and collected products from these facilities for testing. The FDA traced ingredients of the products to identify potential sources of contamination. When available, the products were collected from the homes of case patients and tested for STEC. The implicated company also conducted product testing and shared STEC isolates with the FDA.

FDA laboratory personnel used the primary enrichment methods outlined in the Bacteriological Analytical Manual to test foods for the outbreak strain. The recovery of bacteria was assisted through immunomagnetic separation with the use of the Invitrogen Dynabead Max EPEC/VTEC O121 Kit. The Applied Biosystems Prepman Ultra Kit was used to prepare a template for real-time polymerase chain reaction (PCR). When a sample was found to be positive for STEC by means of real-time PCR assay, the immunomagnetic separation–concentrated pellet was plated to a series of up to six isolation agars (selective, differential, or both) and incubated at 37°C for 18 to 24 hours. Colonies were screened with the Abraxis E. coli Latex Agglutination O121. Isolates were confirmed as E. coli with the use of a Vitek GN microbial identification test card, and real-time PCR was repeated to verify the toxin profile. Isolates were subjected to molecular serotyping by means of Bio-Plex analysis, which was also used to determine the presence of virulence markers. Isolates that were confirmed as E. coli serogroup O121 underwent PFGE and whole-genome sequencing analysis.

**RESULTS**

**CASE IDENTIFICATION**

A total of 56 cases were identified in 24 states (Figs. 1 and 2); 55 were infections with STEC O121, and 1 was an infection with STEC O26. Of the 55 STEC O121 isolates, 40 underwent whole-genome sequencing analysis and were found to be closely related genetically (a difference of 0 to 2 SNPs) (Fig. 3). These 40 isolates were all positive for the gene stx2a; the lone O26 strain was positive for stx1a only. All sequenced isolates were positive for eaeA, and all but 3 of the O121 isolates were also positive for ehIA. Case patients ranged in age from 1 to 95 years (median, 18). A total of 43 of 56 case patients (77%) were female. Sixteen case patients were hospitalized. The hemolytic–uremic syndrome developed in one adolescent girl (infected with STEC O121 that was positive for stx2a, eaeA, and ehIA), but she recovered. No deaths were reported.

**HYPOTHESIS GENERATION**

Leafy green vegetables were commonly reported as having been eaten by early case patients, but the initial questionnaire did not identify any food exposures among case patients that were eaten at a significantly higher frequency than in the FoodNet Population survey. Open-ended telephone interviews then were conducted with 10 patients, all of whom stated that they baked frequently or regularly consumed home-baked foods. Five of the patients recalled baking during the week before illness onset, and 3 others reported that
they might have baked during that period. Of the 5 case patients who remembered baking, 4 reported eating or tasting homemade batter or dough, 3 of whom used brand A flour. The fourth used either brand A or another brand. Two of the patients (a resident of Colorado and a resident of Washington) still had the bags of brand A flour that they had used in the week before illness onset.

Shortly thereafter, state investigators identified 3 ill children who had been exposed to raw flour at restaurants in Maryland, Virginia, and Texas. Restaurant staff had given them raw dough to play with while they waited for their food to be served.

**CASE–CASE ANALYSIS**

The case–case questionnaire included questions about baking, flour, and raw-dough exposures and about other food exposures that had been reported during hypothesis generation. Of the 56 case patients, 33 (59%) in the outbreak completed this questionnaire, as did 84 comparison patients with non-STEC illness. Among these patients, there were 22 matched sets that each contained one STEC outbreak case and one or more comparison cases. Univariable matched analysis showed that STEC infection was significantly associated with baking (odds ratio, 8.79; 95% confidence interval [CI], 2.39 to 32.24) and with the use of brand A flour (odds ratio, 21.04; 95% CI, 4.69 to 94.37), as well as with tasting uncooked or unbaked homemade dough.

![Figure 1. Number of Case Patients, According to State of Residence.](image1)

![Figure 2. Number of Case Patients, According to Week of Illness Onset (December 21, 2015, through September 5, 2016).](image2)
0–2 SNPs (Outbreak clade)

58–130 SNP difference from outbreak clade

58–70 SNPs
72–89 SNPs
69–88 SNPs
72–91 SNPs
110–130 SNPs
E. coli Infections Associated with Flour

or batter (odds ratio, 36.02; 95% CI, 4.63 to 280.17), irrespective of brand, and eating chocolate chips (odds ratio, 15.03; 95% CI, 3.31 to 68.36). Using brand A flour and eating chocolate chips also were found to be significant exposures in the multivariable model (Table 1). Several brands of chocolate chips were reported, however, which made a common source less likely.

PRODUCT TRACE-BACK INVESTIGATION AND TESTING

Trace-back investigation of the two bags of brand A flour collected from patients in Colorado and Washington revealed that the flour from Colorado was unbleached all-purpose flour manufactured on November 14, 2015, and the flour from Washington was bleached all-purpose flour manufactured on November 15, 2015. The two bags were produced in the same facility. The flour that was used in the raw dough given to the children exposed in the Maryland, Virginia, and Texas restaurants also was from this facility, as was flour from three additional bags collected from case patients residing in Arizona, California, and Oklahoma.

Initial testing of the flour collected from the homes of case patients did not identify STEC O121. After additional screening of colonies that are not typical of E. coli, STEC O121 with delayed lactose fermentation was recovered from the Colorado flour sample. The laboratory protocol was modified and used to test subsequent samples. The delayed lactose fermentation observed in the isolates from flour samples was consistent with what was reported for the isolates from clinical samples associated with this outbreak. In total, testing isolated the STEC O121 outbreak strains from flour samples collected from case patients in Arizona, Colorado, Indiana, Michigan,

Table 1. Selected Exposures among Case Patients in the Shiga Toxin–Producing Escherichia coli (STEC) Infection Outbreak and Comparison Patients with Non-STEC Illness.*

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Unmatched Patients</th>
<th>Matched Sets of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case Patients</td>
<td>Patients with Non-STEC Illness</td>
</tr>
<tr>
<td></td>
<td>no./total no. (%)</td>
<td>(%)</td>
</tr>
<tr>
<td>Baked or made homemade cookies, muffins, pancakes, cakes, or other foods containing flour</td>
<td>22/26 (85)</td>
<td>19/77 (25)</td>
</tr>
<tr>
<td>Used brand A flour to make something homemade or from scratch</td>
<td>19/30 (63)</td>
<td>7/78 (9)</td>
</tr>
<tr>
<td>Ate, tasted, or licked any uncooked or unbaked homemade dough or batter</td>
<td>17/30 (57)</td>
<td>3/80 (4)</td>
</tr>
<tr>
<td>Ate any chocolate chips or chunks by themselves or in homemade foods</td>
<td>15/24 (62)</td>
<td>8/76 (11)</td>
</tr>
<tr>
<td>Ate any peanut butter</td>
<td>23/29 (79)</td>
<td>35/77 (45)</td>
</tr>
</tbody>
</table>

* The questionnaire also included questions about exposure to five additional brands of flour, as well as five brands of baking mix. Exposure to each of these brands among case patients did not differ significantly from that among patients with non-STEC illness, and therefore these data were excluded from this table.
† Variables were selected to compare the most probable sources of the outbreak.
‡ Shown are the odds ratios for the selected exposures among outbreak case patients versus patients with non-STEC illness, matched according to state of residence, sex, and age group.

Figure 3 (facing page). Phylogenetic Tree of Selected Shiga Toxin–Producing Escherichia coli Serogroup O121 Isolates Involved in the Outbreak.

The tree is based on 40 clinical isolates (case) and 9 isolates from flour samples (flour). The source of each isolate (type of sample and state abbreviation) is provided after the identification number of each isolate. The numbers at the tree nodes are bootstrap values that indicate the confidence in the clustering on repeated analysis of random subsets of the data (the closer the value is to 100, the higher the confidence is in the clustering). Additional details are provided in the Supplementary Appendix. SNP denotes single-nucleotide polymorphism.
and Oklahoma; of these five samples, four were in original brand A packaging, whereas the Michigan sample was reportedly brand A but was not in original packaging. The isolates from Arizona, Colorado, Michigan, and Oklahoma underwent whole-genome sequencing analysis and were found to be closely related genetically to 40 STEC O121 isolates from clinical samples that represented three of the four outbreak O121 PFGE pattern combinations (a difference of 0 to 2 SNPs) (Fig. 3, and Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org).

FDA inspectors did not identify a source of contamination at the implicated facility, which suggested that the ingredients might have been contaminated further back in the production chain. Company A, the parent company of brand A, also isolated STEC from flour produced at that facility and shared the isolates with the FDA. The FDA conducted whole-genome sequencing on these isolates and identified one STEC O26 strain that was closely related genetically (a difference of 3 SNPs) to one clinical isolate that had previously not been considered a part of the outbreak. This case was subsequently included in the case count on the basis of genetic relatedness and additional epidemiologic information collected.

**PRODUCT RECALL**

This investigation identified flour produced at a single facility as the source of the outbreak. In response, company A issued three recalls of multiple brands of flour produced at this facility. Additional product recalls were issued by other companies that had used the recalled flour in their own products. In total, nearly 250 products containing flour were recalled. Information about these products is provided on the CDC and FDA outbreak Web pages.15,16

**DISCUSSION**

In this investigation, raw flour was identified as the source of an outbreak of STEC infections. In addition to STEC, other foodborne pathogens, including salmonella, have been detected in raw flour and implicated in outbreak investigations, which suggests that, although it is a low-moisture food, flour is a possible vehicle for foodborne pathogens and a potential outbreak source.3,17,18

Linking this outbreak to flour was challenging. Consumption of raw or undercooked flour is not included on most routine state and national foodborne disease questionnaires, so epidemiologists were not initially able to assess whether case patients had consumed raw flour. In addition, many case patients also reported exposure to chocolate chips, but additional epidemiologic and laboratory evidence supporting flour as the source helped to rule out this food. These case patients were baking with both chocolate chips and flour when they were exposed. Some case patients did not report exposure to flour in the week before illness onset, but this is not uncommon in outbreak investigations. Interviews often occur weeks to months after the illness, which makes it difficult for the patient to recall exposures accurately. Another challenge was the fact that most case patients had discarded their flour packaging; therefore, information regarding the production lot, which could have been used to determine the manufacturing location, was often not available. Moreover, the trace-back investigation could not determine whether the implicated flour shared a common source of wheat, because wheat from several states was used to produce the flour and grains from different fields are frequently commingled shortly after harvest and further mixed during transport and milling.

The laboratory component of the investigation also faced difficulties. Laboratory personnel needed to use immunomagnetic-separation techniques to concentrate the pathogen cells in order to isolate STEC from the leftover flour provided by patients. They also used modified screening criteria to isolate the STEC O121 strain, which had delayed lactose fermentation, an unusual characteristic for STEC. Investigations of future outbreaks will need to account for the fact that laboratory procedures using lactose fermentation as a screening step for STEC O121 may reduce the likelihood of recovering the pathogen. This investigation also provided additional evidence that clusters of illnesses with distinct PFGE patterns can be closely related genetically and caused by a common source.19 Apparent differences according to PFGE pattern among isolates that are determined by whole-genome sequencing to be closely related probably resulted from the exclusion of mobile genetic elements (e.g., plasmids and phages) from the whole-genome sequencing analysis because they are less evolutionarily informative.
Although the epidemiologic, trace-back, and laboratory components of the investigation confirmed flour produced in a single facility as the source of the outbreak, the source of the contamination was never identified. On the basis of what is known about the ingredients of flour, wheat is the ingredient most likely to be contaminated, perhaps in the field before harvest. Some farmers use manure from cattle, a reservoir of STEC, to fertilize their wheat fields, which could lead to contamination of the wheat if the cattle are colonized.20 Another source might be white-tailed deer, which are ubiquitous in the United States and are also reservoirs for STEC.21

Given that a specific wheat field was not implicated in this investigation, we could not evaluate whether animal intrusion was a source of contamination.

Since this outbreak resulted in the hospitalization of more than a quarter of case patients and in the development of the hemolytic–uremic syndrome in one, it serves as a reminder of the substantial health consequences of STEC infections. The investigation also highlighted a number of issues that contributed to this multistate outbreak. STEC O121 was introduced into a commercially distributed product at a concentration sufficient to cause a substantial number of illnesses. However, the behaviors of both consumers and retailers increased the risk of illnesses resulting from the contaminated flour. These behaviors included the consumption of raw or undercooked homemade dough or batter, which has long been discouraged because of the known risk of salmonellosis from consuming raw eggs, as well as allowing children to play with raw dough in restaurants and using flour to make play-dough for children at home. Our data show that although it is a low-moisture food, raw flour can be a vehicle for foodborne pathogens.

This article is based on a public health response by local, state, and federal government staff. It did not receive any outside financial support. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the Food and Drug Administration.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank all the outbreak investigation team members in jurisdictions affected by the outbreak who contributed to this investigation, including local and state partners in Alabama, Arizona, Arkansas, California, Colorado, Illinois, Indiana, Iowa, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Montana, Nebraska, New York, Oklahoma, Oregon, Pennsylvania, Tennessee, Texas, Virginia, Washington, and Wisconsin, as well as partners at the FDA and CDC.

REFERENCES


14. GenomeTrakr network. Silver Spring, MD: Food and Drug Administration (http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm363134.htm).


16. FDA investigates multistate outbreak of *Shiga* toxin-producing *E. coli* infections linked to flour. Silver Spring, MD: Food and Drug Administration, September 29, 2016 (http://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm504192.htm).


Copyright © 2017 Massachusetts Medical Society.