August 2014

FSIS Compliance Guideline for Establishments Sampling Beef Trimmings for Shiga Toxin-Producing *Escherichia coli* (STEC) Organisms or Virulence Markers

Table of Contents & Section Links:

I.	**Introduction and Background Discussion**	Page 3
II.	**High Event Periods**	Page 5
III.	**Suggested High Event Period Numerical Criteria**	Page 12
IV.	**Action in a High Event Period**	Page 15
V.	**General Guidance for Verification Testing for STEC**	Page 17
VI.	**Designing Sampling Plans for Verifying Control of STEC**	Page 24
VII.	**Factors Affecting the Design of Sampling**	Page 29
VIII.	**Further Discussion**	Page 31
IX.	**Appendix 1: Detailed Explanation of FSIS's HEP Criteria**	Page 37
X.	**Appendix 2: FSIS Response to Comments**	Page 40

I. Introduction

Food Safety and Inspection Service (FSIS) developed this compliance guideline to assist beef slaughter/fabrication establishments that perform testing for Shiga toxin-producing *Escherichia coli* (STEC) organisms (or virulence markers) using the N60 sample collection method on beef manufacturing trimmings. FSIS requires that establishments perform ongoing verification activities to ensure their food safety system is functioning as intended (9 CFR 417.4(2)) and support decisions made in their hazard analysis (9 CFR 417.2 and 417.5(a)(1)).

Establishment verification testing results on trimmings are likely the best available objective information a slaughter establishment can use to determine the ongoing effectiveness of its slaughter/dressing operation. Establishments can use the information to support decisions made in their Hazard Analysis and Critical Control Points (HACCP) systems. FSIS recommends that establishments incorporate statistical process control procedures into their testing programs. FSIS recommends that establishments use the test results to assess the effectiveness of their controls for preventing contamination during the slaughter operation and verify that they are reducing STEC to a non-detectable level. Establishment sampling programs can be supplemented with other types of verification activities associated with production of other raw ground beef and patty components.

This guidance has been updated to reflect the Agency's recent policy developments. Since FSIS issued this compliance guideline for comments in May 2012, FSIS began testing beef manufacturing trimmings for six non-O157 STECs (O26, O45, O103, O111, O121, and O145) in addition to *E. coli* O157:H7. These six non-O157 STECs are capable of producing Shiga toxin (stx) and intimin (eae). FSIS declared these six non-O157 STECs adulterants in raw, non-intact beef products and product components. (http://www.fsis.usda.gov/wps/wcm/connect/6aa26172-2d27-4534-99d4-8c528b285fd2/2010-0023.pdf?MOD=AJPERES).

The high event period (HEP) guidance provided in this document applies mainly to beef slaughter/fabrication establishments that manufacture 50,000 pounds or more of trimmings daily. Such establishments are likely to conduct sufficient verification testing on same source materials to be able to determine whether a HEP occurred. This guidance includes some general discussion at the end of Chapter III regarding how smaller establishments may choose to define a HEP. This document also includes general information on verification testing, designing sampling plans, and factors affecting the design of sampling. It also includes examples of sampling methods. These topics are covered in Chapters V through VIII and will apply to establishments of any size. This document also provides general information for non-slaughter establishments that produce or receive trimmings, although non-slaughter establishments will not know if problems with

slaughter and dressing procedures have contributed to a HEP situation unless the supplier provides that information as part of a purchase specification program arrangement.

This guideline reflects comments received on the Agency's Draft Compliance Guideline for Establishments Sampling Beef Trimmings for Shiga Toxin-Producing *Escherichia coli* (STEC) Organisms or Virulence Markers issued in May 2012. FSIS has also added an appendix that provides additional information on how FSIS developed its HEP criteria.

Finally, this guidance represents current FSIS thinking. It is considered usable now. FSIS will update the guideline as needed to reflect the most current information available to FSIS and stakeholders. This document provides recommendations rather than regulatory requirements.

This Compliance Guideline follows the procedures for guidance documents in the Office of Management and Budget's (OMB) "Final Bulletin for Agency Good Guidance Practices" (GGP).

More information on the bulletin can be found on the FSIS Web page: http://www.fsis.usda.gov/wps/wcm/connect/fsis-content/internet/footer/policies-and-links/significant-guidance-documents/significant-guidance).

II. High Event Periods (HEP)

HEP are periods in which slaughter establishments experience a high rate of positive results for STEC (or virulence markers) in trim samples from production lots containing the same source materials. That is, the trim was produced from one or more carcasses slaughtered and dressed consecutively or intermittently within a defined period of time (e.g., shift).
A HEP may mean that a systemic breakdown of the slaughter dressing operation has occurred and has created an insanitary condition applicable to all parts of the beef carcass (e.g., primal cuts in addition to the beef manufacturing trimmings and other raw ground beef and patty components). FSIS recommends that establishments identify HEP criteria so that they can determine whether they need to withhold product from commerce when a HEP has occurred because the presence of a HEP may indicate more widespread adulteration of product, beyond the product found positive. If establishments identify and respond to HEP, they will minimize the chance that they release adulterated product into commerce.

KEY QUESTION

Question: What is a HEP?

Answer: High event periods (HEP) are periods in which slaughter establishments experience a high rate of positive results for STEC (or virulence markers) in trim samples from production lots containing the same source materials. A HEP situation may mean that a systemic breakdown of the slaughter dressing operation has occurred and has created an insanitary condition applicable to all parts of the beef carcass (e.g., primal cuts in addition to the beef manufacturing trimmings and other raw ground beef and patty components). This is because HEP indicate a more widespread adulteration of product, beyond any product found positive for the pathogen. By following this guidance and withholding adulterated product from commerce during HEP, establishments are more likely able to avoid costly recalls.

This revised guidance recommends two distinct HEP situation criteria: one type for a localized out-of-control situation, and a second type for a systemic break-down situation. In both situations, FSIS believes that establishments should be concerned if their sampling of trimmings produce a positive rate statistically significantly greater than 5%, rather than 1.5 percent positive, as discussed in the 2008 draft guidance. In such cases the processor should review process control measures and intervention measures used during slaughter, dressing, fabrication, and grinding. The two types of HEP that may indicate out-of-control situations are:

1. A HEP that indicates a *localized* out-of-control event in which some specific occurrence or event causes a clustering of STEC contamination in product.

2. A HEP that indicates a *systemic* break-down or inherent weakness of the process or food safety system.

In both situations, FSIS believes that if an establishment's sampling of trimmings during a period produced a positive rate statistically significantly greater than 5%, there was a severe loss of process control during that period. In such cases, the processor needs to thoroughly review its process control and intervention measures used during slaughter, dressing, fabrication, and grinding.

One difference between a systemic break-down and a localized out-of-control situation is the amount of product that should be assessed to determine whether it may be adulterated. A localized out-of-control situation may affect only the production of one lot, while a systemic break-down may affect more products. Also, a localized HEP may indicate an isolated problem (such as improper application of an antimicrobial on one lot); a systemic HEP may indicate a broader problem (systemic failure to prevent cross contamination among carcasses or improper application of antimicrobial on many lots).

When either of these trigger criteria is reached, the establishment may determine that production lots of beef manufacturing trimmings containing same source materials that were sampled, tested, and found negative should be considered as having a false negative result, depending on the reason for the HEP. If the establishment makes that determination, such product should be diverted to a full lethality treatment or otherwise destroyed.

To develop recommendations for identifying a HEP, FSIS examined industry data collected in 2010 by FSIS inspection personnel from the top 33 slaughter establishments, based on production volume (heads slaughtered). Of the 33 establishments, 32 responses were received, 19 had clear definitions of a HEP, 2 had clear but incomplete definitions because they did not specify a time frame (which FSIS interpreted to be a day), 10 had unclear definitions of a HEP, and 1 did not have a definition. Of the 21establishments that had clear definitions (including the two FSIS interpreted), 7 were using a 5% threshold definition; there were 3 that had definitions greater than 10%.

KEY QUESTION

Question: Why should establishments evaluate their test results using HEP criteria?

Answer: When establishments experience a HEP, FSIS believes with a high degree of confidence of poor processing or poor food safety controls. Therefore, establishments need to take extraordinary action to ensure that adulterated product does not enter commerce.

Based on these results, FSIS selected a target of 5%.
That is, FSIS would consider an establishment's process to be within a HEP if the percent positive within a set time (1 day, shift, etc.) is 5% or greater.

FSIS would not consider the establishment's process within a high event period if the percent positive is less than 5% within a given time period.
FSIS developed its HEP criteria using a very high degree of statistical confidence. That way, FSIS has a very high degree of confidence that the process's percent positive truly exceeds 5%. FSIS did not select a lower target because it did not want to define HEP criteria that would be as, or more, rigorous than those of a large number of establishments. FSIS did not select a higher target (e.g., 10%) because the Agency believes such a target could result in many cases where poor processing, as defined by most of the industry, would not be detected as a HEP.

The HEP criteria FSIS has provided in this guidance indicate exceptional events of poor processing. FSIS would consider an establishment's process to be within a HEP if, during the period of sampling, the (true) process percentage of positive is not less than 5%. To avoid incorrectly saying a HEP occurred, FSIS required a specified degree of statistical confidence before asserting a HEP occurred. More detailed information on how FSIS developed its HEP criteria can be found in Appendix 1: Detailed Explanation of FSIS's HEP Criteria.

Industry typically makes decisions to identify HEP based on presumptive positive results (or what FSIS terms potential positive results) from initial screening tests that produce a high percentage of false positives.[1] Percentages of presumptive positive results are greater than percentages of confirmed results. FSIS identified a target that was in accordance with today's present industry standard.

Most establishment testing methods include an enrichment step followed by differential screening specific to STEC or their virulence markers.
Positive results during screening tests require further testing to detect STEC.
If the establishment does not perform further testing, it should treat positive screen results as confirmed positives. FSIS considers those results positive for STEC.

FSIS recognizes that many establishments test for other STEC or their associated virulence markers and treat those positive screen results as positive for STEC. Establishments can apply the guidance in this document to such positive screen results. Therefore, this document refers to *E. coli* O157:H7 and the six non-O157 STECs as "STEC" and positive screening tests for STEC virulence markers.

The 2008 guidance specified that an *E. coli* O157:H7 percent positive of greater than 1.5%, for samples collected using the N60 collection method, indicated that

[1] The FSIS analysis for the presence of *E coli* O157:H7 has two screening stages before confirmation. See MLG 5A.02 (10/01/10). Also see: http://www.fsis.usda.gov/wps/wcm/connect/316b1e53-beba-4f2b-971a-b587fe4744e2/Mlg_5A_02.pdf?MOD=AJPERES for a flow chart of FSIS's procedures.

a process may be out of control and thus in a HEP situation. Based on that percent positive, the prior guidance recommended that four positive *E. coli* O157:H7 results out of 91 consecutive N60 samples should be seen as indicating a loss of control.

FSIS changed the recommended target from the value given in its 2008 draft guidance of 1.5% positive to 5% positive. FSIS made this change for two primary reasons. First, FSIS recognizes that many establishments treat a potential or presumptive positive sample result as if the sample was confirmed to contain viable *E. coli* O157:H7. These practices result in a higher positive rate than that seen for FSIS verification testing that is specifically for viable *E. coli* O157:H7 that have been confirmed positive. Second, FSIS made this change to

KEY NOTE

As noted earlier, beef slaughter/fabrication establishments that manufacture 50,000 pounds or more of trimmings daily are likely to conduct sufficient verification testing on same source materials to be able to determine whether a HEP occurred based on these criteria.

a higher target value to increase confidence that an insanitary condition likely occurred during the slaughter/dressing operation. With a greater target value that the process is statistically significantly greater than 5%, the establishment and the Agency would have greater confidence that the food safety system is truly out of control for many situations with an identifiable cause than compared to its confidence when using a target of 1.5%. FSIS does not expect such HEP situations to happen often during any 12-month period when an establishment's slaughter dressing operation is properly functioning. FSIS did not want to recommend HEP criteria for establishments unless it can be highly confident that there are identifiable factors that contributed to the high percentage of found positive results.

Establishments may choose to use the earlier guidance based on 1.5%. As is discussed below, establishments may choose to develop stricter HEP criteria than FSIS is recommending. By choosing the stricter HEP criteria, the establishment reduces its vulnerability for releasing product into commerce that could test positive at a subsequent point in processing or that could be associated with illness. Establishments should be able to support whatever HEP criteria they use.

HEP criteria are most useful to establishments that have rigorous testing programs. FSIS recommends that slaughter/fabrication establishments conduct sampling and testing of trim at a frequency sufficient to find evidence of contamination surviving the slaughter and dressing operation (optimally every production lot) in an effort to ensure that adulterated product does not enter

commerce. Microbiological testing is necessary because not all contamination can be seen with the naked eye. It is also important to detect all contamination because high levels of contamination may significantly reduce the safety margin afforded by antimicrobial treatments.

Establishments grouping five combo bins of trimmings into production lots represented by one N60 sample may be less capable of discerning a HEP situation than establishments that collect an N60 sample from a one combo bin production lot. However, establishments that group multiple combo bins into a production lot may have a scientific basis for selecting samples or grouping samples that allows them to identify a HEP effectively; they may have had a contract study conducted for them based on their own in-plant conditions that supports their lotting practices and shows that their sampling and testing has a high probability of detecting positives when present. In addition, establishments that exclude exterior fatty trimmings from calculation of a HEP situation may be less capable of identifying a HEP situation.

Additional assistance and information on these matters can be found in publications of the Beef Industry Food Safety Council, at: http://www.bifsco.org/groundbeef.aspx

For the purpose of this document there are two types of a HEP that may indicate out-of-control situations:

1. A HEP that indicates a *localized* out-of-control event in which some specific occurrence or event causes a clustering of STEC contamination in product.

2. A HEP that indicates a *systemic* break-down or inherent weakness of the process or food safety system.

FSIS recommends that establishments evaluate their testing results for both short term (local) and long term (systemic) periods for which the positive rate is substantially greater than that expected or typically observed within production days or shifts. Establishments can evaluate their processes during both periods by applying a set of criteria within different moving windows of testing results. If the establishment exceeds one of the HEP criteria, FSIS believes there is a high degree of confidence that particular events occurred that indicates poor processing or poor food safety controls.

Below are criteria establishments may use for determining whether they have experienced a HEP.

1. For a local HEP: 3 or more STEC (or virulence markers) positive results out of 10 consecutive samples from production lots containing same source materials; that is, the trim was produced from one or more

carcasses slaughtered and dressed consecutively or intermittently within a defined period of time (e.g., shift); and

2. For a systemic HEP:

    a. 7 or more STEC (or virulence markers) positive results out of 30 consecutive samples from production lots containing same source materials.

    b. Establishments that test more than 60 samples per day from production lots containing same source materials can use Table 1 below for determining criteria for a systemic HEP.

KEY QUESTION

Question: How did FSIS develop its HEP criteria?

Answer: FSIS developed the HEP criteria based on a high degree of statistical confidence that the establishment exceeded 5% positive. For the local HEP criteria, FSIS used close to 99% (98.95%) confidence. For the systemic HEP guidance, FSIS used about 99.95%. FSIS expects that establishments will improve their percent positives over time, and, therefore, FSIS plans to adjust the HEP criteria accordingly. See Appendix 1 for more details.

Table 1 provides, for a given number of samples (window) the number of positive results that would provide close to 99% confidence that the establishment's process percentage of positive results during the period of sampling exceeds 5%. The criteria in Table 1 apply for local HEP and for systemic HEP that apply for a day (or shift) of processing (2b above). The numbers of positive results within these windows, if obtained, provide high degrees of confidence that the establishment has poor processing or poor food safety controls. The moving window method of monitoring process control is a simple and useful tracking procedure.

FSIS also is recommending a HEP sample criterion based on obtaining 7 or more positive results within any 30 consecutive samples from production lots. This criterion provides a greater degree of confidence that the process percentage of positive results exceeded 5% than those for the other criteria. FSIS chose a greater degree of confidence (about 99.95%) because in this situation FSIS believes that the establishments should examine rigorously not only trim product but also other primal and subprimal products. For the other HEP criteria, Table 1 provides numbers that would result in a 99% confidence that the 5% target would have been exceeded.

Table 1 also includes the observed percentage of positive samples. The reason the observed percentage of positive results is greater than the 5% target FSIS used to establish HEP is because the HEP criteria are based on a very high degree of statistical confidence that the process's percentage truly exceeds the

5% target. Establishments may use this table for their testing programs to determine if they have a HEP, or may develop their own.

Table 1: HEP Criteria when Establishment test more than 60 samples per Day or local HEP for 10 consecutive samples

Unacceptable # Positives	Number of Samples	Confidence	Observed Percentage of Positive
3	10	98.8%	30.0%
8	61	98.9%	13.1%
9	74	98.9%	12.2%
10	86	98.9%	11.6%
11	100	98.9%	11.0%
12	113	98.9%	10.6%
13	127	98.9%	10.2%
14	141	98.9%	9.9%
15	155	98.9%	9.7%
16	169	98.9%	9.5%
17	184	98.9%	9.2%
18	198	98.9%	9.1%
19	213	98.9%	8.9%
20	228	98.9%	8.8%

FSIS is not providing a tolerance for an acceptable number of STEC (or virulence markers) positives. Rather, FSIS is providing guidance on when the number of positive results within a certain number of samples indicates a HEP occurrence. In such situations, negative-tested production lots are possibly contaminated because they were likely produced under insanitary conditions. The establishment would need to determine whether the lots are releasable.

KEY QUESTION

Question: Is FSIS establishing a tolerance for an acceptable number of STEC (or virulence markers) positives?

Answer: No, FSIS is providing criteria that indicate exceptional events of poor processing and require extraordinary action so that adulterated product does not enter commerce.

The establishment–specific process positive rate may differ from the rate used to construct the above example. These rates may differ depending on the time of year and increase during high prevalence seasons. Consequently, a specified positive rate for a given establishment at a given time should be identified by indicating that a different positive rate was being achieved consistently and product has low likelihood of being adulterated. However, FSIS would consider an establishment to experience a HEP if it experiences a HEP according to FSIS's criteria regardless

of the time of year. Further, deviations from previously obtained positive rates should be construed as presumptive evidence that the process is out of control and would warrant investigation to find and eliminate any potential causes for the positive results. As part of their supporting documentation for their hazard analysis, FSIS recommends that establishments document their criteria for identifying a HEP.

One example for how establishments might develop their own criteria would be to determine an upper bound (limit) process percent positive and then determine how many actual sample results they will use to show whether they have exceeded that upper bound[2]. FSIS expects that slaughter/fabrication establishments are subjecting 100 percent of production lots of trim to N60 verification testing. In addition, FSIS expects that establishments would have a more rigorous verification testing program during the high prevalence season (from spring into mid-autumn) in order to have greater confidence that increased contamination is not passing through the slaughter/dressing operation into the trim production lots. More rigorous verification testing programs might include more restrictive HEP criteria. In addition, small establishments or those that produce product infrequently might choose a different set of criteria from those provided by FSIS.

III. Sample HEP Numerical Criteria

The end of Chapter III below has specific suggestions.
The following tables are provided to help establishments derive parameters for determining whether they have experienced a HEP. The tables provide specified numbers of positive results (first column) occurring within a specified number of samples (entries within the remaining columns) from production lots.

Those indicate that the true percent positive of STEC findings (or virulence markers) would be greater than or equal to the specified percent positive given in the column headings, for the following percent confidence intervals:

[2] There are other tracking procedures, such as calculating and graphing cumulative sums of differences of results from a specified target (CUSUM) and exponentially weighted moving averages (EWMA), which do not require such determinations.

- with 95 percent confidence (Table 2);
- close to 99 percent confidence (Table 3); and
- close to 99.95 percent confidence (Table 4).

The latter two tables show how FSIS developed the criteria for localized or systemic HEP. In the tables below, the test result from one composite sample of multiple slices (e.g., N60 sample) is considered one positive or negative result.

Table 2: Lower Bounds of Percent Positive, Based on Number of Samples Tested
True positive percent of STEC (or virulence markers) findings is greater than corresponding lower bound percentage in column with 95% confidence, given the number of positive results (rows) within corresponding number of samples (interior table entries)

Number Positive	0.50%	0.68%	0.75%	1.0%	1.5%	2.0%	3.0%	3.5%	5.0%
2	71	52	47	35	24	18	12	10	7
3	164	120	109	82	55	41	27	23	16
4	274	201	182	137	91	69	46	39	28
5	395	290	263	198	132	99	66	57	40
6	523	385	349	262	175	131	88	75	53
7	658	484	439	329	220	165	110	95	67
8	797	586	532	399	266	200	134	115	81
9	940	692	627	471	314	236	158	135	95
10	1086	799	725	544	363	273	182	156	110
11	1235	909	824	618	413	310	207	178	125

Based on Table 2, if there were 4 or more positive results within 69 samples, then there would be 95% confidence that the process positive percent exceeds 2%.

Table 3: Lower Bounds of Percent Positive, Based on Number of Samples Tested

True positive percent of STEC (or virulence markers) findings is greater than corresponding lower bound percentage in column with about 98.85% confidence, given the number of positive results (rows) within corresponding number of samples (interior table entries)

Number Positive	0.50%	0.68%	0.75%	1.0%	1.5%	2.0%	3.0%	3.5%	5.0%
2	32	23	21	16	11	8	5	5	3
3	92	68	62	46	31	23	16	13	10
4	172	127	115	86	58	44	29	25	18
5	266	196	178	133	89	67	45	39	27
6	369	272	247	185	124	93	62	54	38
7	481	354	321	241	161	121	81	70	49
8	598	440	399	300	200	151	101	87	61
9	720	530	481	361	241	181	121	104	74
10	846	623	565	424	283	213	143	123	86
11	976	718	652	489	327	246	164	141	100

Table 4: Lower Bounds of Percent Positive, Based on Number of Samples Tested

True positive percent of STEC (or virulence markers) findings is greater than corresponding lower bound percentage in column with about 99.95% confidence, given the number of positive results (rows) within corresponding number of samples (interior table entries).

Number Positive	0.50%	0.68%	0.75%	1.0%	1.5%	2.0%	3.0%	3.5%	5.0%
3	32	24	21	16	11	8	6	5	4
4	75	55	50	38	25	19	13	11	8
5	132	97	88	66	45	34	23	20	14
6	200	148	134	101	68	51	35	30	21
7	278	205	186	140	94	71	48	41	30
8	363	268	243	183	123	92	62	54	38
9	455	335	304	229	153	116	78	67	48
10	552	407	369	277	186	140	94	81	58
11	654	482	437	329	220	166	111	96	68

Based on Table 4, if 5 positive results occur within the set of 20 samples, then there is about 99.95% confidence that the positive percent exceeds 3.5%. The establishment may decide that if its percent positive exceeds 3.5%, then the establishment has experienced a HEP.

Establishments might have reason to collect a number of samples representing product processed under similar conditions and thus is indicative of the processing during a set period. For example, an establishment might run product

during a given time, or for a given day (shift), and take 20 samples of that product. In that case, the tables or similar calculations used for creating the tables can be used for deciding how many positive results within sets of 20 samples would indicate a percent positive greater than that expected.

Small establishments that test infrequently might decide to develop other criteria for determining whether they have experienced a HEP. For example, a small slaughter establishment may test 5 samples and find 2 of them positive. For a small establishment that does not test frequently, two positive results (or even 1) might indicate a lack of control in the production of that product and thus could be considered as a HEP.

If Tables 2-4 above do not meet an establishment's needs for determining high event criteria appropriate for the establishment, the establishment should contact *askFSIS* at http://askfsis.custhelp.com/ and categorize its question as "Sampling" within the system. Through the *askFSIS* system, establishments can obtain expert advice on the design of HEP criteria.

IV. Action in a High Event Period

IV. Action in a High Event Period (HEP)

In a robust testing program, negative results normally indicate that product may be released in commerce. However, during a HEP, the establishment needs to consider whether negative-tested lots of trimmings are releasable, and whether primal and subprimal product produced from the same source materials as the trimmings may be positive for STEC.

The actions taken in response to a HEP could depend upon the findings of the investigation of the positive results. If slaughter establishments experience a HEP, they should assess what happened during the slaughter and dressing process and take appropriate action that would ensure only unadulterated product is released into commerce. Studies have shown that STEC is present in the hides and intestinal contents of cattle and, therefore, can contaminate the surface of the carcass, trimmings, ground beef, and

KEY QUESTION

Question: What actions does FSIS recommend in response to HEP?

Answer: When a HEP occurs, the establishment needs to consider whether negative-tested lots of trimmings are releasable, and whether primal and subprimal product produced from the same source materials as the trimmings may be positive for STEC. For a local HEP, establishments may not need to sample, test and hold primals and subprimals. However, during systemic HEP, primal and subprimal cuts should be sampled and tested even if treated with an antimicrobial treatment. FSIS recommends that establishments test food contact surfaces for STEC (or virulence markers), and if the surface is found positive consider product that came into contact with those surfaces to be adulterated.

other beef products (e.g., primals, subprimals, and mechanically tenderized or enhanced beef) during slaughter, fabrication, grinding, and processing.

The process of removing the hide and intestinal tract requires care, and even under good manufacturing practices, occasional contamination of the carcass meat will occur from direct contact of the hide to the carcass, contact of the hide to equipment, hand-to-hide-to-carcass contact, aerosolization when removing the hide, or puncture of the intestinal tract. Slaughter and dressing procedures should be designed to minimize, to the maximum extent practical, cross-contamination of carcasses with the contaminants from the hide and intestinal tract.

As FSIS stated in FSIS PHIS Directive 6410.1, Verifying Sanitary Dressing and Process Control Procedures by Off-Line Inspection Program Personnel (IPP) in Slaughter Operations of Cattle of any Age, the Agency expects that establishments will slaughter and process cattle in a manner designed to prevent contamination from occurring at any step in the process and will use decontamination and antimicrobial intervention treatments as necessary to address any contamination that may result from the implementation of the slaughter process or otherwise occur on the carcasses.

If a slaughter establishment believes a HEP has occurred, FSIS recommends that the processor review process control measures and intervention measures used during slaughter, dressing, fabrication, and grinding. Such controls may include measures to reduce the pathogen load on incoming animals, measures to ensure that contamination of the carcass does not occur during slaughter or dressing procedures, decontamination or antimicrobial treatments, and measures to minimize carcass-to-carcass contact and cross contamination. Ensuring and verifying that such controls are indeed working is crucial to preventing future HEP.

The actions taken in response to an out-of-control signal could depend upon the findings of the investigation of the positive results. If the establishment finds the cause for the HEP and takes corrective action to prevent positive results from recurring, then an increase in the sampling rate would not be needed. However, the establishment needs to have a high degree of confidence that the corrective actions will be effective before reducing the intensity of its testing. Until such a high degree of confidence is obtained, FSIS recommends that the establishment conduct increased testing when it experiences a HEP. For example, the establishment could increase sampling rates by either defining smaller lots of trimmings (1 combo bin instead of 5 combo bins) or selecting additional samples from the 5 combo bin lots.

During systemic HEP, FSIS recommends that primal and subprimal cuts be sampled and tested, even if treated with an antimicrobial treatment. In addition, during systemic HEP, FSIS recommends that establishments test food contact

surfaces for the presence of STEC (or virulence markers). If they detect the pathogen, establishments should consider product that came into contact with those surfaces to be adulterated.

These recommendations are not regulatory requirements. However, by taking these additional steps, establishments will be able to better ensure that they do not release adulterated product into commerce. Therefore, these additional steps may reduce the likelihood of costly recalls. During local HEP, FSIS recognizes that establishments may determine that less product may be affected or implicated by the positive results than in systemic HEP. Establishments may not need to sample, test, or hold primals and subprimals during local HEP.

The prevalence of STEC has been greater in cattle coming to slaughter during the warmer months (from spring into mid-autumn – the "high prevalence season") than the colder months. Thus, HEP should be especially anticipated during the high prevalence season. Extra steps should be implemented to increase confidence that contaminated product is not released into commerce for use in raw beef during the high prevalence season compared to the low prevalence season. Such steps could include more frequent monitoring and verification of both slaughter and dressing procedures, additional antimicrobial reduction treatments, or sampling and testing additional product. FSIS also recommends increasing sampling and verification testing during the high prevalence season.

During traceback activities, Enforcement, Investigations, and Analysis Officers (EIAOs) will gather information about the production of the product including the use of anti-microbials, prevention of cross-contamination, sanitary conditions, and relevant purchase specifications. Furthermore, as part of their traceback investigations, EIAOs will review slaughter establishment test results to determine whether the establishment has experienced a high event period (HEP).

Establishments are free to develop their own HEP definition as an alternative based on their unique operations. If establishments define HEP differently than FSIS, establishments should support their definition of HEP and provide the information to FSIS during traceback. For purposes of FSIS traceback activities, FSIS will identify HEP events based on the establishment's HEP criteria, provided the establishment's criteria is appropriately supported. During FSIS traceback activities, FSIS will evaluate the establishment's definition and support for defining HEP and determine whether the establishment has taken all affected product into account. If the establishment has not developed its own HEP criteria or its criteria is not supported, EIAOs will determine whether the establishment experienced a HEP based using the guidelines provided in this document. In the event the establishment has not developed or appropriated supported HEP criteria, the specific HEP criteria FSIS will use during traceback are:

1. For a local HEP: 3 or more STEC (or virulence markers) positive results out of 10 consecutive samples from production lots containing same source materials; that is, the trim was produced from one or more carcasses slaughtered and dressed consecutively or intermittently within a defined period of time (e.g., shift); and

2. For a systemic HEP: 7 or more STEC (or virulence markers) positive results out of 30 consecutive samples from production lots containing same source materials.

Based on the results of their traceback activities, EIAOs will make recommendations whether regulatory and enforcement actions are warranted. The District Manager will then determine whether adulterated product entered commerce; if it has, whether to contact the FSIS Recall Management and Technical Analysis Staff; and whether enforcement actions are appropriate. When FSIS requests that establishments recall product, FSIS looks at several factors to determine the scope of a recall, including the establishment's processing and sanitation procedures, and whether there is any finished product reincorporated into fresh product (rework).

V. General Guidance for Verification Testing of STEC

Why is STEC Verification Testing Important?

KEY QUESTION

Question: Why does FSIS recommend that establishments test beef trimmings for STEC?

Answer: Establishment verification testing results on trimmings are likely the best available objective information a slaughter establishment can use to determine the effectiveness of its slaughter/dressing operation.

Because microbial contamination is not visible to the naked eye, microbiological testing is needed to verify that the slaughter and dressing procedures that are designed to prevent microbial contamination are effective. Also, an establishment may incur considerable expense if it becomes necessary to recall contaminated product from commerce[3]. This action becomes necessary when trimmings that have been subjected to antimicrobial interventions are later found positive for STEC or when other production lots from same source materials – i.e., fabricated from a single, common source rather than multiple, commingled sources – are found positive. For these reasons, robust sampling and testing programs that can find product containing STEC can be highly cost-effective.

[3] FDA has estimated that a recall can cost government and industry $3-5 million. ("Preliminary Regulatory Impact Analysis and Initial Regulatory Flexibility Analysis of the Proposed Rules to Ensure the Safety of Juice and Juice Products" (63 FR 24258; May 1, 1998). The cost covers manufacturer, retailers and State, local, and Federal authorities.

Extensive sampling of trimmings and careful evaluation of test results can help establishments identify areas of poor processing for corrective action.
FSIS recommends that establishments continually strive to decrease STEC (or virulence markers) percent positives. FSIS expects establishments that investigate and correct problems will improve processes and decrease percent positives over time. Consequently, FSIS recommends that both slaughter establishments and receiving establishments test source product, including trimmings, for STEC (or virulence markers). FSIS recommends testing finished product even if the source material has been tested and found negative.
The reason for this recommendation is that negative test results on samples of product do not imply that product is free of STEC for the following reasons: there may have been pockets of contamination in the product that were not in the actual sample tested at the slaughter establishment, the product might have

KEY POINT

FSIS recommends that both slaughter establishments and receiving establishments test source product, including trimmings, for STEC (or virulence markers).

become contaminated after it was sampled at the slaughter establishment, or STEC (or virulence markers) cells within the actual sample tested might not have been detected at the slaughter establishment because their numbers at the time of testing were below the limit of detection. If a receiving establishment finds incoming product intended for grinding or other raw, non-intact use positive for STEC (or positive in a screening test but not confirmed negative), that product is adulterated, although it may be treated to eliminate the pathogen. The receiving establishment should inform the supplier of the positive test results.

For What Organism Should Establishments Test?

It is useful to conduct verification testing for associated organisms that include STEC (e.g., a screen methodology for pathogenic STEC) and maintain records of results. Measurements of ubiquitous organisms such as *Enterobacteriaceae*, aerobic plate counts (APC),[4] or generic *E. coli* can be used to evaluate the effectiveness of process controls designed to limit or eliminate microbial contamination. Frequent measurement of APC may capture a short-term trend, which would be useful for quality control, both before and after the sanitary dressing processes. However, such measurements, while helpful for ensuring

[4] Measuring the level of APC on pre-eviscerated carcasses might be useful for evaluating the effectiveness of a sampling program and antimicrobial interventions (see T. A. Arthur, et al., 2004, J Food Protection 67(4): 958-665).

microbial process control, cannot be used as a substitute for determining the actual presence or absence of STEC in the final product.

The decontamination and antimicrobial treatments applied during the slaughter and dressing operation should be designed to remove, to the maximum extent practical, contamination with pathogens. Each establishment should know the limits of capability of its slaughter and dressing operation for reducing microbial contamination as evidenced by objective data, such as for APC or other indicator organisms of process control on the carcass immediately after hide removal, before the establishment applies any antimicrobial interventions.

KEY QUESTION

Question: What is the significance of non-O157 STEC?

Answer: The Centers for Disease Control and Prevention (CDC) estimates that there are approximately 175,905 domestically acquired foodborne illnesses associated with all Shiga toxin-producing *E. coli* (STEC) annually (Scallan et al, 2011)[5]. *E. coli* O157:H7 is the most well known STEC and, according to the CDC, is annually responsible for approximately 63,153 (36%) of the domestically acquired foodborne STEC illnesses. The remainder of the illnesses associated with STEC (112,752 or 64%) are caused by non-O157 STEC. While more than 50 non-O157 STEC serogroups have been associated with human illness, 70-80 % of confirmed non-O157 STEC illnesses are caused by six STEC serogroups – O26, O45, O103, O111, O121, and O145. These illnesses can be equivalent in severity to those caused by *E. coli* O157:H7. In the U.S, at least one outbreak and several sporadic illnesses from non-O157 STEC serogroups have been associated with ground beef products.

Many establishments that produce raw, non-intact beef products, such as ground beef, incorporate antimicrobial interventions such as organic acid sprays in their processing. These methods should be effective in controlling non-O157 STEC. However, many firms will want to implement their own testing programs.
A prudent establishment would use a test method that includes all hypothetical strains of *E. coli* O157:H7 and the target non-O157 STEC, either typical or variant organisms with these STECS serotypes, that would be identified using FSIS confirmatory testing procedures and criteria and that increases the likelihood of detecting low level contamination by these pathogens.

FSIS recognizes that industry uses non-cultural methods that detect alternative target analytes for STEC including, but not limited to, eae and stx.
Establishments may increase the likelihood of detecting all hypothetical strains and low levels of contamination by these pathogens in a variety of ways, including but not limited to using a test method that also is used by a regulatory body or that is validated and certified by an independent body (e.g.., AOAC, AFNOR, MicroVal, or NordVal). An establishment may also opt to use a test method that is subjected to a robust validation using the FSIS cultural method as a reference.

[5] Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, Jones JL, and Griffin PM. 2011. Foodborne illness acquired in the United States – major pathogens. Emerg Infect Dis. 17(1):7-15.

KEY QUESTION

Question: Are test kits available to test for STEC?

Answer: Several companies have developed or are developing test kits to detect at least the six relevant STEC serogroups. Some kits have been submitted for review by validation bodies. In addition, some kits have been submitted for FSIS review and have received "letters of no objection" from the Agency. FSIS developed guidance for evaluating test kit performance and uses this criteria for evaluating test kits that have been submitted to FSIS for review. FSIS provides a summary table that describes the non-O157 STEC test methods that FSIS has received and reviewed and for which it has issued letters of no objection. The summary table is available on FSIS's website. FSIS continues to review other test kits submitted for review. For establishment testing or testing on behalf of an establishment, FSIS recognizes that other criteria, while not used specifically by FSIS for identification of a non-O157 STEC, may be a significant and expedient indicator of the presence of non-O157 STEC in products. Such tests might be applied as rapid screening procedures to expedite analyses. If an establishment uses or contracts with a laboratory that uses such rapid screening procedures, and product is found positive by that test, the regulations require the establishment to take appropriate corrective action and to ensure the proper disposition of adulterated products following a positive test result (9 CFR 417.3). The establishment will need to define and support the criteria it uses to define the sampled lot.

How Frequently Should Establishments Collect a Sample?

Sampling and testing of trimmings for STEC (or virulence markers) should occur at a frequency sufficient to find evidence of contamination (e.g., pathogens) surviving the slaughter and dressing operation. Optimally, every production lot should be sampled and tested before

KEY QUESTION

Question: How often does FSIS recommend that slaughter establishments sample and test trim to verify their HACCP system?

Answer: FSIS recommends that slaughter/fabrication establishments conduct sampling and testing of trim at a frequency sufficient to find evidence of contamination (e.g., pathogens) surviving the slaughter and dressing operation (optimally every production lot) in an effort to verify and ensure that adulterated product does not enter commerce.

leaving the slaughter establishment and again before use at the receiver. Establishments that do not slaughter but do produce trimmings should report their test results back to the slaughter supplier in order for the supplier to assess the adequacy of its slaughter and dressing practices, as well as antimicrobial treatment programs. Through this feedback, an investigation of the possible reasons for the contamination getting through the slaughter and dressing operation can be conducted and could lead to the identification and correction of possible deficiencies.

During the high prevalence season months (from spring into mid-autumn), the frequency of such testing should be increased compared to that of the other months in order to have increased confidence that contamination is not affecting the food safety system.

Chapter VI of this document includes information on designing sampling plans. Defining a sampling plan involves establishing the procedures the establishment will use, including how it will collect a sample, the size of the units it will collect, the number of samples it will collect, the frequency with which it will collect samples, and the procedure it will use to analyze samples.

What Corrective Actions Should Establishments Take in Response to Positive Test Results?

Under 9 CFR 417.3, establishments are required to identify corrective actions in response to every deviation from a critical limit or a deviation not covered by specified corrective action. An STEC positive would fall into one of these two categories that require corrective actions. Corrective actions required in the regulations include identifying and eliminating the cause of the deviation (if STEC are addressed in the HACCP plan) or reassessing the HACCP plan and determining whether changes to it are necessary (if STEC is not addressed in the HACCP plan).

Process control of STEC can be evaluated by tracking past sample results, enabling establishments to tell the difference between an occasional, sporadic, positive result and a loss of process control as indicated by many positive results over time. If past sample results lead establishment management to believe the process is out of control, the establishment should carefully investigate to find all contributing causes. This type of investigation would be more involved than a follow-up investigation when an occasional positive result is found. The finding of an out-of-control process may implicate product in other production lots produced during the period that the process was out of control or from the same source material.

How are Corrective Actions Different in Response to a HEP?

It is important to note that a HEP situation (localized and systemic) likely means that insanitary conditions occurred during the slaughter/dressing operation such that contamination is widespread across production lots. When a HEP situation occurs, negative test results from the production lots of trimmings made from the same source materials as trimmings found positive during the HEP may not be reliable. [6] Therefore, those production lots that tested negative may not be microbiologically independent of those directly associated with the HEP. In other words, even though an N60 product sample from a production lot tested negative, trimmings produced from the same source materials as the production lots directly associated with the HEP are also potentially contaminated.

KEY QUESTION

Question: What does the occurrence of a HEP mean?

Answer: A HEP likely means that insanitary conditions occurred during the slaughter/dressing operation such that contamination is widespread across production lots. When a HEP situation occurs, establishments should take appropriate precautionary steps to ensure adulterated product is not released into commerce. The establishment specifically needs to consider whether negative-tested lots of trimmings are affected and whether intact primal and subprimal product produced from the same source materials as the trimmings may be positive for STEC.

When a HEP occurs, establishments should take appropriate precautionary steps to ensure adulterated lots of raw beef are not released into commerce. The establishment specifically needs to consider whether negative-tested lots of trimmings are affected and whether intact primal and subprimal product produced from the same source materials as the trimmings may be positive for STEC.

Generally, if primals are not commingled before packaging, and the establishment prevents cross contamination among primals, primals can be considered independent lots. Normally, FSIS does not typically consider primal cuts designated for intact use to be adulterated if contaminated with STEC. During a HEP, however, unless the establishment has controls in place to ensure that the primals are not used for non-intact purposes, such primals may be considered adulterated because they were prepared under insanitary conditions. Establishments that subject primals to an antimicrobial treatment as part of a routine production process may be able to demonstrate

[6] The statement does not imply that the usual N60 sampling and testing is not reliable regarding their sensitivity and specificity. However, it is reasonable to assume that during the HEP, the incidence and levels of contamination could be greater than normal. Negative results during a HEP could be resulting from clustered contamination within a lot. Thus, a negative result may not mean that the lot is actually negative. To suggest re-testing product that already tested negative in light of an HEP just means that FSIS is asking establishments to take action to decrease the assumed increased risk that exist in light of the HEP.

that the primals are not adulterated, provided they have on-going verification testing results to affirm that contamination was not evident.

It should be noted that a recent large-scale recall of beef that included primal cuts was associated with a HEP (http://www.fsis.usda.gov/wps/portal/fsis/topics/recalls-and-public-health-alerts/recall-case-archive/recall-case-archive-2009/!ut/p/a1/jZDBCoJAEIafpQdYdlZF9CgLppa7SGS2lxjEdMFUTDz09CmdDKVmTjN8Px8zVNGMqgZHXeKg2wbreVb2DRKwmcshkr7nQyhMP3XEnoG0J-C6AFw2A2kiD5yDI8w_8xvlwa989IfA6GMel1R1OFREN_eWZn2RY12THJ8FwT6v9FisLokB4NILVUsNsKknzckKImGCtL6BlT98gO1Du8c5ex0D0KG3ewMRQGX5/?1dmy&current=true&urile=wcm%3apath%3a%2Ffsis-archives-content%2Finternet%2Fmain%2Ftopics%2Frecalls-and-public-health-alerts%2Frecall-case-archive%2Farchives%2Fct_index286a). Although all or most trim was diverted to cooking, including trim that tested negative for *E. coli* O157:H7 (or other STEC organisms or virulence markers), primal cuts that had not been treated with an antimicrobial entered commerce. Illnesses were associated with the trim derived from the untreated primal cuts.

KEY QUESTION

Question: In addition to responding to individual positive test results, what does FSIS recommend that establishments use their test results to determine?

Answer: FSIS recommends that establishments monitor their test results for two distinct HEP criteria: one for a localized out-of-control situation, and a second for a systemic break-down situation. In both situations, establishments should be concerned if their sampling of trimmings produce a positive rate statistically significantly greater than 5%.

VI. Designing Sampling Plans for Verifying Control of STEC

How Should Establishments Define Their Lots?

Designing a sampling plan involves identifying many factors, including among others, the lot size and the amount of product from each lot that is to be sampled and analyzed. Perhaps the most important step in designing a sampling plan is the definition of a lot of product. The results (positive or negative for the presence of STEC or virulence markers) may determine the disposition of the product within the selected lot and possibly other product as well, depending on how the lots are defined.

Trimmings from each supplier should be tested separately. Limiting product in a lot to that from a single supplier could help decrease the extent of product that would be recalled or sent for cooking when a positive test result is obtained. An establishment should be sure to always define the production lot size before sampling. An establishment should not redefine it during testing or after results are known.

Lots should be defined so that if a positive result is found from one lot, the product in other lots is microbiologically independent and is not implicated. FSIS has stated (67 FR 62325; Oct. 7, 2002) that when one lot of trimmings tests positive, lots constructed from the same source material likely would be implicated. FSIS would expect the establishment to have a scientific basis that justifies why any raw ground product produced from those source materials should not be considered to be adulterated (67 FR 62325; Oct. 7, 2002, p. 62333). One way to avoid the results for one lot implicating another is to ensure that the lots are microbiologically independent.

KEY POINT

Lots should be defined so that if a positive result is found from one lot, the product in other lots is microbiologically independent and is not implicated.

The establishment should have a sound basis for defining the lot and is responsible for determining the lot of product represented by the sample.

<u>Suggestions for defining microbiologically independent lots:</u>

1. Product from different carcasses can be considered as independent lots provided the meat from the carcasses was handled so as not to cross-contaminate other carcasses.

2. Defining lots based on microbiological testing would be acceptable if the sample collection method is designed to have a high confidence of detecting positive results when STEC is present in a production lot.

3. Processing interventions that limit or control STEC contamination can help to define the lot.

4. Beef manufacturing trimmings and raw beef components or rework carried over from one production period to another may expand the implicated lot in the event of a positive result.

5. Sanitation Standard Operating Procedures (Sanitation SOP) or any other prerequisite programs used to control the spread of STEC cross-contamination among raw beef components during production can help to define the lot. The following may lead to cross-contamination of raw beef components during production and may expand the implicated lot in the event of a positive result:
    - improper sanitary dressing procedures
    - insanitary product contact surfaces on equipment, such as machinery and employee hand tools
    - improper employee hygiene

What Questions Should Establishments Consider when Designing a Sampling Plan?

In designing a sampling plan, an establishment should consider the following questions.

A. What products are to be tested?

   Trimmings or other source materials that are supplied to grinders, including cheek meat and head meat (see FSIS Directive 10,010.1)

B. The size of the lot: what amount of product (i.e., the lot) is to be represented by a sample?

   The establishment should define how much product is going to be grouped together to constitute a "lot" (e.g., combo bins of trimmings; boxes of packaged head meat or cheek meat).

   Note: FSIS strongly recommends that the lot definition not be redefined. It is unacceptable to change the lot definition based on the results of testing.

C. How is the sample going to be collected?

   1. STEC (or virulence markers), when present, is not evenly distributed throughout a production lot. Therefore, a collection method that selects product at multiple sites within the lot or multiple production intervals within a given lot is more likely to detect pockets of contamination than a sampling plan that samples at fewer sites or production intervals.

   2. For trimmings, potential contaminants will be on the exterior surface of the product that was exposed during the slaughter and dressing process. Therefore, collection methods that provide more surface area for the test increase the sensitivity of the sampling (i.e., many thin slices of the exterior exposed fat and lean tissue).

3. For trimmings, samples can be collected by:

    - Obtaining 60 slices from the exterior surface of product within the lot that are as thin as possible resulting in the desired sample size (grams) to be collected.

    - "Plug" collection, where product is collected by inserting a specially designed "tube" between pieces of meat so as to excise the trim (exterior areas) of adjacent pieces. This procedure is performed many times, by inserting the tube at randomly selected locations, to ensure that a certain minimum number of exterior surface pieces are collected and achieving the proper weight for the sample.

    - Randomly selecting slices of trimmings from trim in combo bins. If the establishment produces several types of trimmings, it should include them all in the sampling program. Trim from exterior surfaces should be prioritized for collection.

    - Core drilling, where product is collected at several places in the combo bins by drilling a hole, approximately 25mm in diameter, into surface of meat through a template. The product is thus extracted through a coring tube and can be taken from fresh or frozen trim.

    - For frozen trimmings, using a sanitized band saw at 12 points around the edges of a 60-pound frozen block. To make up N60, samples should be collected from five randomly selected frozen blocks.

    With all these collection methods, specifications should be designed to ensure that a high percentage of the collected product that is to be used for testing consists of exterior surface tissue.

KEY POINT

With all collection methods, specifications should be designed to ensure that a high percentage of the collected product that is to be used for testing consists of exterior surface tissue.

D. How much of the collected product is analyzed?

1. FSIS recommends that the entire sample be analyzed. To accommodate laboratory testing methods that limit the amount of material per analysis, subsamples could be formed and each subsample analyzed in the laboratory. Thus, multiple analyses may be needed. Not analyzing the

entire sample could lead to a significant increase in false negative results (negative results found when the product is actually positive) compared to when the entire amount is analyzed, so that results could be misleading. Laboratory methods used should be effective in detecting the pathogen.

NOTE: A sampling plan using the N60 collection method and analyzing a 325-375 gram composite sample means that the weight of each of the 60 slices that is 'represented' in the tested material needs to be about 6.25 grams (375 grams/60 slices = 6.25 grams per slice).

E. How frequently should establishments test?

Optimally, every production lot should be sampled and tested.

F. How effective is the testing method?

1. FSIS recommends that the establishment understand and have written documentation regarding how the laboratory is testing the sample, in regard to the size of the sample analyzed and the analytical method that is used.

2. FSIS recommends that laboratory methods be "fit for purpose" and ensure detection of very low levels of STEC (or virulence markers) that may have survived antimicrobial treatments. FSIS recommends that methods be approved or used by a recognized government or independent body (e.g., FSIS, FDA, AOAC, AFNOR, ISO).

3. In some circumstances, multiple samples may be "pooled" after enrichment to save costs for testing. Because negative broths can dilute positive broths in the pooled test broth, "wet-pooling" analytical methods should ensure that sensitivity is not compromised. Wet-pooling refers to combining multiple samples for a single screening test after the samples have been enriched, i.e., incubated overnight in a broth as the first stage for detecting a pathogen.

4. It is important for testing laboratories to follow the testing protocol as written to ensure the method will perform as expected. This includes pre-warming the enrichment broth to the incubation temperature before incubation to help ensure the greatest sensitivity, particularly for methods using enrichment periods less than 15 hours.

KEY QUESTION

Question: Why is it important to analyze the entire sample?

Answer: Not analyzing the entire sample could lead to a significant increase in false negative results (negative results when the product is actually positive), which increase the likelihood that adulterated product enters commerce.

5. In circumstances when a test result for pooled samples is positive, it may be appropriate to re-test the individual sample-specific enrichments in an attempt to identify contaminated product more accurately. In such a procedure, it is important that the storage of the enrichments does not cause a decrease in the sensitivity of the individual test as compared to the pooled test.

SUMMARY: KEY ELEMENTS OF A SAMPLE PLAN

A sampling plan used to verify process controls should address the following:

1. Products to be tested
2. Lot size (usually in pounds and number of combo bins)
3. Statistical sampling method for selecting lots; percentage of lots that are sampled (lot sample)
4. Slice size (dimensions) and number of slices that comprise a sample
5. Collection method for selecting samples and slices from a selected lot
6. Procedures for preparing a sample for analysis (See Chapter VIII)
7. (Sub) sample size analyzed in a laboratory
8. Laboratory testing methods used (including sample size analyzed, enrichment procedures and size of portions analyzed)
9. Actions to take when samples are positive (See Chapter V)

KEY POINT

It is unacceptable to change the lot definition based on the results of testing.

VII. Factors Affecting the Design of Sampling

A critical limiting factor in a sampling plan is the maximum sample size that the laboratory can analyze. Given this maximum sample size, the sample is characterized by the number of slices and the slice size. Because the contamination occurs on the surface of the meat, slices should be as thin as possible and focus on surface tissue. Because it is expected that STEC organisms (or virulence markers) when present would be distributed unevenly in clumps, in constructing samples it is advisable to use many small sample slices rather than few larger slices (all slices should be of the same thickness). Using many small slices provides a more "representative" sample of the lot and greater likelihood of finding contamination. However, the limiting factor here is the time to collect many slices. At present, an N60 sample involves collecting 60 slices of a specified dimension. An N120 sample with slices ½ the surface area of those used for N60 would be expected to provide a greater likelihood of finding positive results, given everything else being equal. However, the time needed to collect an N120 sample might be twice as long as needed to collect an N60 sample. With limited resources, a likely consequence of the longer time needed to collect an N120 sample would be that fewer lots or combo bins would be sampled, thus losing the advantage of N120 sampling over N60 sampling. Over the years, the N60 sample has become the standard sample for beef trim products. It is important to remember that changing the slice size of samples or the number of slices for a sample could have an impact on the expected percentage of positive findings.

Several factors can guide establishments in designing their sampling plans. Two of them are discussed here: percentage of positive samples in the product and degree of confidence desired for a given sample to test positive.

A. Percentage of positive samples of STEC (or virulence markers) in the product

- The percentage of positive samples is determined as the number of positive samples for the pathogen divided by the total number of samples tested, multiplied by 100. The process percent positive is the expected percentage of positive samples over time.

B. Degree of confidence desired for a given sample to test positive

- The distribution of cells of STEC (or virulence markers) will depend on the levels on the carcasses and effectiveness of the control measures used by the establishment during slaughter, dressing, and fabrication (e.g., intervention treatments, temperature, and sanitation). An establishment that has verified that its control measures (e.g., organic acid spray wash or control of incoming materials) are effective in reducing contamination by the pathogen should have lower levels of STEC (or virulence markers).

- The percentage of contaminated slices within a contaminated lot might likely be small. Thus, large numbers of slices for a sample are needed to determine with high confidence that a sampled lot has STEC (or virulence markers) cells. Table 5 shows the number of slices that would need to be collected to have 95% confidence of detecting STEC (or virulence markers) in the sample consisting of a random collection of n slices, assuming a specified true percentage of contaminated slices within the lot.

Table 5: Calculations used to derive the number of slices given in the table assume that the 'sizes' of the slices are the same (based on slice size used for N60 samples).

Percentage positive slices	0.5%	1%	1.7%	2.5%	5%	7.5%	10%	15%	23%
Number of slices needed	598	299	178	119	59	39	29	19	12

Table 5 shows that about 60 selected slices are needed to have a 95% confidence that contamination will be detected when the percentage of potential slices within a single lot (available for selection) that are contaminated is equal to 5%. Selecting 12 slices only provides the same degree of confidence of finding a positive when the true percentage of contamination is about 23% within the lot.

The above table suggests that if sensitivity greater than that of N60 sampling is desired, more slices (i.e., more surface area) would be needed. Since each slice varies in thickness, and thus in weight, the entire N60 sample is portioned into a 325g or 375g analytical portion size, which is a requirement of the FSIS method and the portion size industry typically uses, respectively. It is possible to obtain more sensitivity by taking larger (number of slices) samples. For example, two N60 samples per lot could be collected, for a total of 120 slices (of the same size). From Table 5, this would provide about 95% confidence of detecting contamination if 2.5% of the slices within the lot were contaminated. The costs with such sampling, however, could be double that of N60 sampling, assuming that all lots were to be tested, because the time to collect the samples could be doubled, and two samples rather than one sample would be analyzed. To help mitigate the latter cost, the wet-pooled procedure for testing could be used.

VIII. Further Discussion

A collection method known as N60 (mentioned above) is often used for monitoring incidence of STEC (or virulence markers) in beef trim products manufactured by the industry. The '60' refers to the number of slices that are used in constructing the composite sample. The slices are collected randomly from the lot in order to help ensure a good 'representative' sample from the product within the lot.

KEY QUESTION

Question: Why is it important to take samples originating from the exterior carcass tissue?

Answer: Contamination introduced during the slaughter process occurs on the surfaces of the carcass. So collecting samples from the original surfaces of the carcass increases the likelihood of detecting contamination if it is present.

The collection method may be as follows:

Lot size: 5 combo bins consisting of 2,000 pounds each, for a total of 10,000 pounds trim

Number of slices: 60 slices of product sliced from the surface of the meat, 12 slices from each combo bin

Slice size: Each slice is about 6.25 grams and 1/8 inch thickness

Sample size: 375 grams, composited from the 60 slices

1. Take 12 slices of product, randomly selected from and throughout each combo bin, such that each consists of product of about 6.25 grams with thickness of no more than 1/8 inch, to help ensure that the sample will consist of as many slices from the carcass surface (where the STEC organisms or virulence markers are more likely to reside) as feasible to achieve the desired sample weight (325 or 375 g). As a guide, the dimensions of the sample can be about 3 inches in length and 1 inch in width.

KEY QUESTION

Question: Why is it important that sample slices be thin?

Answer: Sample slices should be thin. This maximizes the number of slices that can be collected to achieve the desired sample weight (325 or 375g).

2. If for some reason, there are less than 5 combo bins from which product is to be collected, a total of 60 surface slices from the available combo bins would still be taken. For example, if there are 2 combo bins to be used for grinding, 30 surface slices from each combo bin to make a total of 60 surface slices would be taken; if there were 3 combo bins, 20 slices, and so forth, would be taken.

3. Combine (composite) slices for every lot – the combined 60 slices is referred to as a composite sample.

4. Store the sample at temperatures between 7 and 10 ºC (44 - 50 °F), and send to the laboratory. The sample should be analyzed within 24 hours of collection.

5. At the laboratory the sample must be mixed before selecting the material to be analyzed. It is important that an approximately equal amount of material from every slice be included in the material that is being analyzed.

6. At the laboratory, if necessary, create sub-samples to be analyzed separately (typically five 75-gram sub-samples), though some procedures allow for the whole 375 gram sample to be analyzed.

7. Incubate (enrich) each sub-sample to ensure adequate growth of any STEC cells.

8. Analyze each sub-sample for the presence of STEC – confirm as positive or negative all presumptive positive results for STEC or assume presumptive results are positive.

9. Investigate possible sources of the contamination, the process, and the controls that have been designed to prevent contamination if a result is positive.

10. Dispose of the positive lot and all other implicated product or send for full lethality.

The method of analysis should be equivalent to that of the current method that the FSIS laboratories use as cited in the Microbiological Laboratory Guidebook (MLG) (http://www.fsis.usda.gov/wps/portal/fsis/topics/science/laboratories-and-procedures/guidebooks-and-methods/microbiology-laboratory-guidebook).

Some Variations of N60:

The N60 collection method being used by most establishments involves 5 combo bins defining a lot. This method was designed to detect contamination slice-specific incidence at a within-lot contamination of 5%; lower percentages, averaged over the lot, would not be detected so readily. The results from Table 5 indicate a possible reason for this, namely the number of slices per combo bin (12 for N60) is too small to detect contaminated combo bins.
Consequently, FSIS recommends establishments decrease the production lot size from 5 combo bins to 1 combo bin in order to provide greater assurance that contamination is detected within combo bins. That is, for the N60 method, as described above, for each combo bin there would be 60 surface slices collected. If the combo bin-specific test is positive, the product in the combo bin is sent for cooking; if negative, the product from the combo bin is sent for grinding, provided

that there is no evidence that the process is out of control, based on the percent of positive results for neighboring lots that had been tested, or for any other reason known that could permit contamination not to be removed as effectively as normal.

Wet-Pooling of Samples

Using one combo bin as a lot may increase the cost of analysis. One way to help reduce the laboratory costs of analyses when testing each combo bin would be to enrich each N60 sample, and then pool aliquots of the individual enrichments from the five sampled combo bins. This means that the pooled aliquot represents five N60 samples. This lab method should ensure that a single positive sample pooled with multiple negative samples does not compromise the sensitivity of the testing method (the sensitivity of the test compared to the test used by FSIS).

Diagram 1: The variation of N60 sampling shown (below) is an example of a wet pooled sample. In this situation, if the laboratory pooled sample is positive, then the laboratory would separately analyze the 5 enriched samples (each representing an N60 sample from each of the combo bins) to ascertain which of the combo bins represented in the laboratory pooled sample likely contributed to the positive pooled aliquot sample result. If the enrichment step is done properly, at least one of the 5 enriched samples would be found positive.
The establishment would divert the one combo bin represented by the positive sample to further processing, such as cooking, to destroy the pathogen. In such a procedure, it is important that the storage of the enrichment samples does not cause a decrease in the sensitivity of the individual sample test as compared to the test on the pooled sample.

If none of the individually analyzed N60 enriched samples was found positive, then this might indicate a problem with the enrichment procedure or with the sample handling. In such a case, all product within the 5 combo bins, even though they individually tested negative, would need to be cooked or disposed of because the testing did not identify the positive bin.

Diagram 1: Variation of N60 Sampling

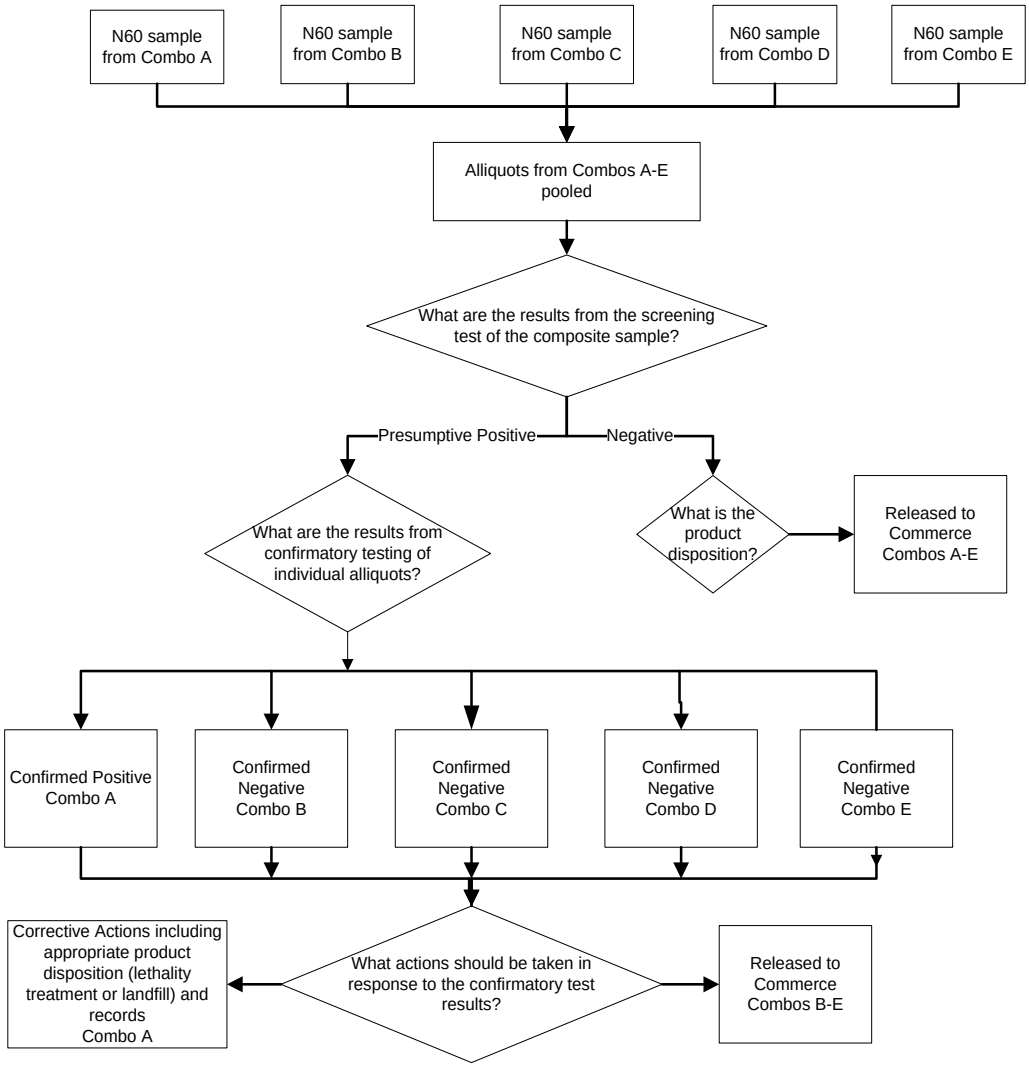
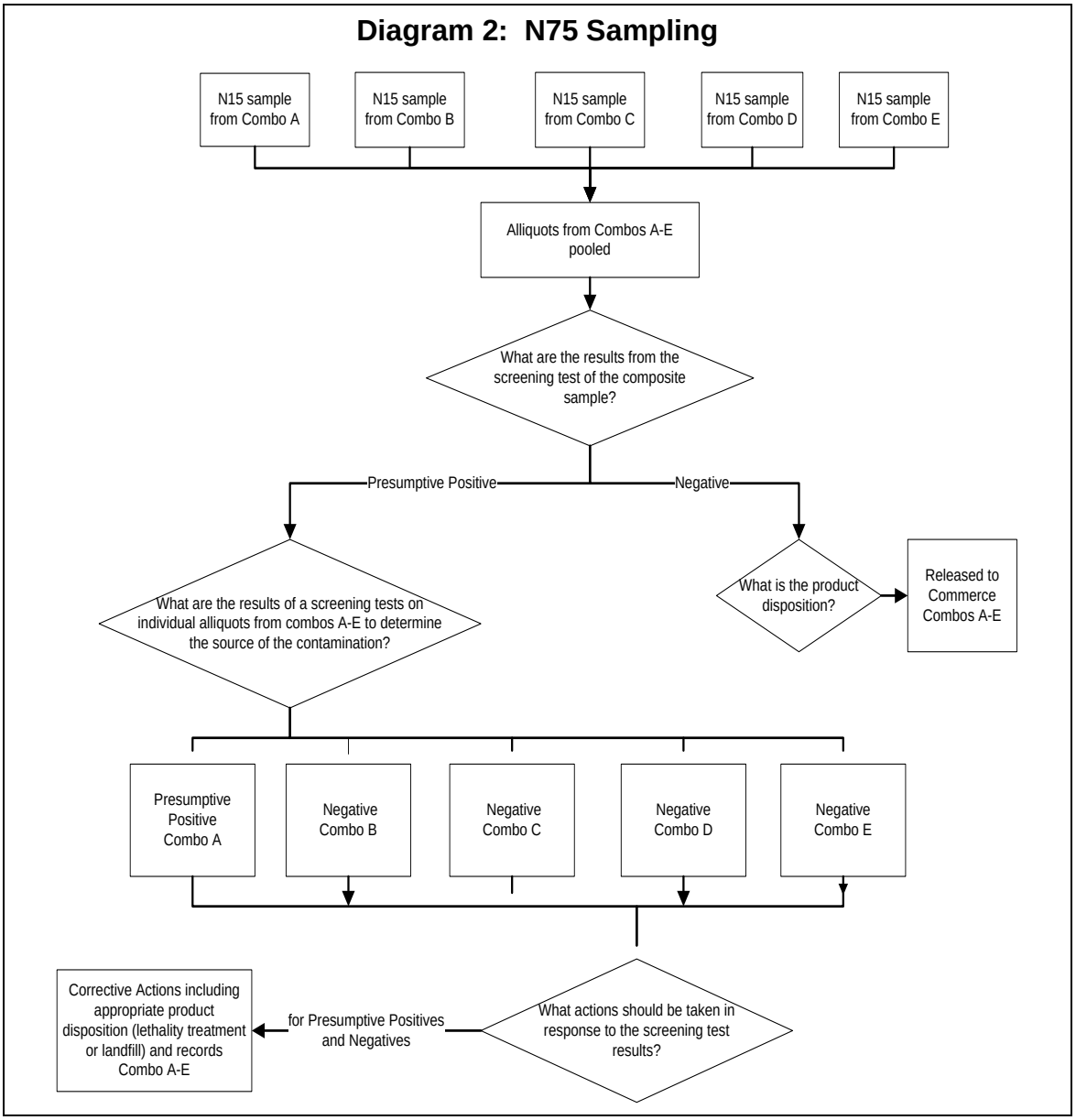


Diagram 2: Variation of N75 Sampling (below) depicts another variation of an STEC testing program with wet pooling. In this example, a lot is defined to be 5 combo bins, with 15 portions from each of 5 combo bins enriched. Aliquots of the enrichment from 5 combo bins are pooled and tested for an initial screening test for STEC. If the screening test for the pooled composite sample is positive, then each of the individual N15 aliquots are tested with screening tests to determine which combo may be the source of contamination. However, in this situation, further screening to determine which of the combo bins may be the source of contamination is not possible because the N15 enrichment broths do

not provide enough statistical confidence to rule out any one of the combo bins as the source of contamination. Therefore, the establishment should treat all 5 combo bins as positive.

Diagram 2: N75 Sampling



IX. Appendix 1: Detailed Explanation of FSIS's HEP Criteria

FSIS assumes that, at any time, there are identifiable processing factors that have an effect on the expected percentage of found positive results. Thus, the expected percentages of found positive results can vary as these identifiable factors vary over time. By defining criteria for high event periods (HEP), FSIS is trying to identify periods of processing such that if the process as constituted during that period were to continue, the expected percentage of found contaminated samples would be greater than the expected percentage that normally is obtained.

There are also factors that influence the results that cannot be identified or specifically control for and, thus, make it impossible to know how many positive results would be found among a fixed number of samples. The impact of these factors FSIS summarizes in a mathematical model by assuming that there is a random component (set of factors) of the process that affects results in a way that FSIS can only describe through a specified probability distribution, as explained in more detail below.

Accordingly, for describing results associated with a processing period, the model identifies two parameter objects: 1) the expected "long term" percentage of found positive results (that would be found over an infinitely large number of tested samples); and 2) a probability distribution that describes the distribution of the number of positive results among a small number of tested samples.
The model's objective is to identify periods of processing (considering these to be High Event Periods) where the first parameter - the expected percentage of found positive results - is greater than some specified value. What makes this difficult is that the evidence is based on small numbers of tested samples and, thus, the results obtained from these samples might not reflect the true expected percentage of positive results associated with the process because of the randomness of the obtained results caused by factors that are not identifiable or uncontrollable. FSIS does not want to recommend declaration of a HEP unless it can be highly confident that there are identifiable factors that contribute to the high percentage of found positive results.

The degree of confidence obtained for a correct inference that the process was out of control as defined here for HEP depends on the number of samples tested and the observed percentage of positive results. To compute the degree of confidence associated with results from a small number of samples, FSIS specified the probability distribution that describes random results. FSIS followed conventional procedures for doing this by assuming the number of positive results, m_t, for n_t samples in a period indexed by t, is distributed as a binomial distribution, $b(m_t, n_t, p_t)$, where p_t is the assumed expected percentage of found positive samples (or the probability that a specific sample will be found positive), and n_t is the number of samples being tested. If, from the results, FSIS statistically infers that p_t is greater than 5%, with a high degree of confidence,

FSIS says that an HEP has occurred. Because of randomness, it is possible that a process during a period that has an expected percentage of positive results - the value of p_t – less than 5% could have, for the few samples collected in that period, an observed percentage of positive results greater than 5%. If, however, the observed percentage of positive samples is sufficiently large, FSIS can statistically infer with high confidence that the p_t is greater than 5%.

Consequently, imposing statistical criteria to provide a high degree of confidence that the (true) process percentage of positive results exceeds 5% means that the observed percentages of positive results that lead to HEPs are greater than the 5% target. The smaller the number of samples, the greater the observed percentage of positive results needed to infer than the true or expected percentage of positive results produced by the process was greater than 5%.

The greater the degree of confidence desired, the greater the observed percentage of positive sample needed, for the same number of samples. Because industry is continuously testing product, FSIS set a relatively high degree of confidence of nearly 99% confidence before declaring a local HEP or a systematic HEP for a day when more than 60 samples were tested. This means that, if the expected percentage of positive results were actually 5%, there is only about a 1% probability that an HEP would be (incorrectly) declared for the period being considered. FSIS believes that when HEPs are identified, there would likely be factors that could be identifiable and controllable as causing the high percentage of positive samples.

Calculations:

FSIS assumed that the number of found positive results, k, in n_t samples is distributed as a binomial distribution as described above. This means the probability of k found positive results in n_t samples assumed an expected percentage of found positive samples, p_t, is given by the formula.

$$P(k|n_t, p_t) = \binom{n_t}{k} p_t^k (1 - p_t)^{(n_t - k)} \qquad (1)$$

The degree of confidence, C, that m_t found positive samples in n_t tested samples suggests that p_t is greater than 5% is calculated as:

$$C(m_t, n_t) = 100\% \sum_{k=0}^{m_t - 1} P(k|n_t, 0.05) \qquad (2)$$

where k is an index for the summation of the probabilities of k found positive results. For example, FSIS specified that in 10 samples, if 3 (or more) are found positive, there would be at least nearly 99% confidence that the expected process-specific percentage is greater than 5%. The more exact confidence is 0.988496. If only 2 positive results were to occur, the confidence would be 0.914 which is considerably less than FSIS's approximate target of nearly 0.99. Accordingly, FSIS set its criterion for 10 samples to be 3 or more positive

findings. Table 1, presented in the main document, provides the degree of confidence obtained for the set of criteria associated with the local HEP (3 or more positive results out of 10 samples) and the systemic HEP when more than 60 samples in a day are tested. Also included is the observed percentage of positive samples.

In summary, FSIS built in a statistical tolerance, requiring nearly 99% confidence before considering that the process, with an observed percentage of positive test results greater than 5%, has a true expected percentage of positive results greater than 5%. In such circumstances, the establishment should take special actions. FSIS also identified a more severe type of HEP (short-term systematic), for which FSIS stated that nearly 99.95% confidence is needed to assert that the true expected percentage of positive results exceeds 5%. In such instances, the establishment should take maximum action. The criterion is 7 or more positive results in 30 samples. The confidence, based on Eq. 2, that the true expected percentage of positive results is greater than 5% when 7 positive results are found in 30 samples is 99.943%, which is near 99.95%.

FSIS using tolerances (i.e., requiring a high degree of confidence) increases the likelihood that time and resources the establishment spends in response to HEP will be fruitful in finding problems that will lead to improvement. Improvement over time would lead to a lower percentage of positive results. FSIS will be monitoring the percentage positive (through its sampling programs), and expects that in the coming years, the target percentage that FSIS will use will decrease from the present 5%.