Risk-based *Listeria monocytogenes* Sampling Data – Data Documentation

Overview

The data covered by this documentation are the sampling results of FSIS' routine risk-based *Listeria monocytogenes* (RLm) sampling program that was designed to detect *Listeria monocytogenes* in the environment of a production establishment. The RLm sampling program collects a set of product, food contact, and environmental (non-food contact) samples at an establishment, which is done in conjunction with a risk-based Public Health Risk Evaluation (PHRE) or a routine Food Safety Assessment (FSA). Additional information can be found on the FSIS Laboratory Sampling Data web page.

Data postings are part of the Agency's efforts to prevent pathogens from entering the food supply throughout the farm-to-fork continuum. Posting these datasets may help industry identify repetitive subtypes and implement control measures. It may allow researchers to identify trends to address basic research questions or to develop new diagnostics or therapies such as vaccines.

These datasets are posted for informational purposes only and are not sufficient to determine if there is an association between multiple samples. Therefore, this data should not be used to identify foodborne illness outbreaks, associate samples with foodborne illness outbreaks or determine whether two or more samples are causally related.

The FSIS number is a unique identifier for retrieving whole genome sequence data from the National Center for Biotechnology and Information (NCBI) Pathogen Detection Isolates Browser. The allele codes included in this dataset provide a convenient naming method for reporting Whole Genome Sequencing (WGS) data. Because allele codes can change over time as more WGS data becomes available, a date stamp allows the data to be used in reports.

FSIS, the Centers for Disease Control and Prevention (CDC) and other public health partners monitor WGS information gathered from samples in real time and use sophisticated epidemiological tools to identify whether the cumulative findings might signal a foodborne illness outbreak. Outbreaks identified through this process are announced to the public through CDC's website and FSIS' Outbreak Response Page. The FSIS investigative process is described in FSIS Directive 8080.3.

Two datasets will be provided: archived and current. The archived dataset will provide data starting from October 1, 2013, up to the final day of the previous fiscal year (FY) of the report execution date (e.g., the data posted in April 2022 was through the end of FY21). The archived dataset will be updated annually. The current dataset will provide data starting from the first day of the FY following the final date available in the archived dataset through the end of the previous fiscal quarter of the report execution date (e.g., data posted in April 2022 was through the end of FY22 Quarter 1). The current dataset will be updated quarterly.

A set of samples collected at an establishment is sent to one of the three FSIS laboratories for analysis. Each row in this data set represents one sample collected.

Isolate characterization data will not be publicly posted in the datasets until the full characterization profile is completed.

Data contained in this dataset on tested product from establishments are not sufficient to determine an association with human illnesses. Further epidemiologic information is needed to determine if there is an association among the non-clinical isolates and human illnesses.

Data Dictionary

- EstablishmentID
  - A unique numeric identifier that is used to identify an establishment across data tables in the FSIS databases.
- EstablishmentNumber
  - A letter/number combination uniquely identifying each establishment.
- EstablishmentName
  - The name of an establishment on its FSIS grant of inspection.
- State
  - The state where the establishment is located.
- ProjectCode
  - A short name given to easily identify an FSIS sampling project.
  - See FSIS Directive 10,240.5, Revision 3 for additional details
  - Projects used in this dataset:
    - RLMPRODC – Routine risk-based product samples taken for post-lethality-exposed product.
    - RLMPRODC_EL – Routine risk-based product samples taken for post-lethality-exposed product that were analyzed at the FSIS Eastern Laboratory.
    - RLMPRODC_MWL – Routine risk-based product samples taken for post-lethality-exposed product that were analyzed at the FSIS Midwestern Laboratory.
    - RLMPRODC_WL – Routine risk-based product samples taken for post-lethality-exposed product that were analyzed at the FSIS Western Laboratory.
    - RLMCONT – Routine risk-based food contact surface swab samples taken for post-lethality-exposed product.
    - RLMCONT_EL - Routine risk-based food contact surface swab samples taken for post-lethality-exposed product that were analyzed at the FSIS Eastern Laboratory.

        - RLMCONT_MWL - Routine risk-based food contact surface swab samples taken for post-lethality-exposed product that were analyzed at the FSIS Midwestern Laboratory.
        - RLMCONT_WL – Routine risk-based food contact surface swab samples taken for post-lethality-exposed product that were analyzed at the FSIS Western Laboratory.
        - RLMENVR – Routine risk-based samples of brine or chill water that does not come into direct contact with post-lethality-exposed product.
        - RLMENVC – Routine risk-based non-food contact environmental surface swabs samples exposed to the post-lethality environment.
        - RLMENVC_EL - Routine risk-based non-food contact environmental surface swabs samples exposed to the post-lethality environment that were analyzed at the FSIS Eastern Laboratory.
        - RLMENVC_MWL - Routine risk-based non-food contact environmental surface swabs samples exposed to the post-lethality environment that were analyzed at the FSIS Midwestern Laboratory.
        - RLMENVC_WL - Routine risk-based non-food contact environmental surface swabs samples exposed to the post-lethality environment that were analyzed at the FSIS Western Laboratory.
- ProjectName
    - The name of the FSIS sampling project.
- FormID
    - The form number used to uniquely identify a specific sample.
- CollectionDate
    - The date the FSIS inspector collected the sample at the FSIS-regulated establishment.
- SampleSource
    - The type of product collected in the sample.
- ProductionAlternative
    - Definition: The production alternative used by the establishment to produce the product that was sampled.
    - Possible Alternatives
        - ALT 1 – The establishment uses a post-lethality treatment (PLT) to reduce or eliminate *Lm* in the product and an antimicrobial agent or process (AMAP) to limit or suppress growth of *Lm* in the product.
        - ALT 2 PLT (Post-Lethality Treatment) – The establishment uses a PLT to reduce or eliminate *Lm* in the product.
        - ALT 2 AMAP (Anti-Microbial Agent or Process) – The establishment uses AMAP to limit or suppress growth of *Lm* in the product.
        - ALT 3 – The establishment relies on sanitation alone to control *Lm* in the processing environment and on the product.
- ListeriaMonocytogenesAnalysis

  - The result of the analysis for *Listeria Monocytogenes* (*Lm*) in the sample.
    - Negative = *Lm* was not found in the sample
    - Positive = *Lm* was found in the sample.
      - All RTE samples with a positive *Lm* result were either prevented from going into commerce due to test and hold or were recalled.
- LmPFGEPattern
  - The specific pattern identified from Pulsed-Field Gel Electrophoresis (PFGE), which is the laboratory technique used to produce a deoxyribonucleic acid (DNA) fingerprint for a group of the same type of bacteria. FSIS discontinued PFGE characterization in March 2019 and replaced the technique with the more modern whole genome sequencing (WGS) characterization.
- LmAlleleCode
  - A code assigned by CDC-PulseNet based on the number of differences in pre-defined genes in the WGS data. The allele code also includes the date when FSIS retrieved the allele code from PulseNet. It is possible for PulseNet to adjust the allele code after it was retrieved. The data format is:
    allele code[space][pipe][space]mm/dd/yyyy (e.g., LMON1.0 – 1.2.3.4.5.6 | 01/01/2021). When PulseNet is unable to assign an allele code, the entry will be: Allele Code Ineligible[space][pipe][space]mm/dd/yyyy (e.g., Allele Code Ineligible | 01/01/2021).
- LmFSISNumber
  - A unique identifier for retrieving Whole Genome Sequencing (WGS) data for a *Lm* isolate from the National Center for Biotechnology and Information (NCBI) Pathogen Detection Isolates Browser. NCBI developed the Browser to help users learn about the sequences they contribute. NCBI has provided a video introduction to this browser, and this document contains a table outlining information available in NCBI'S Pathogen Detection Isolates Brower for additional reference.

Relationship to Other Data

This data can be combined with other FSIS datasets using the EstablishmentID variable.

Notes and Limitations

Information about FSIS sampling laboratories and procedures can be found on the FSIS website on the Laboratories & Procedures web page and the Microbiology Laboratory Guidebook (MLG) web page.

NULL values indicate that the specific variable is not available for that record.

When a sample screens positive for a pathogen, there normally is only one isolate (e.g., subtype determined using serology [serotype], PFGE, antibiotic resistance, allele code, or whole genome sequence) derived from laboratory confirmation procedures. During such procedures, the enrichment broth is streaked on agar plates, and those plates are subsequently examined for typical pathogen colonies. The laboratory staff ordinarily picks no more than one typical isolated colony from any one plate. On very rare occasions, more than one typical colony may be picked for confirmation. In such circumstances, the multiple isolate data (e.g., *Salmonella* serotype) are separated by a semicolon.

Confirmed positive bacterial isolates can become unsuitable for additional testing during shipping, transfer, or storage. This result is final as further characterization cannot be performed. When this occurs, the isolate characterization result fields (e.g., serotype, antimicrobial resistance, allele code) will be populated with "Final result – characterization not available" and the FSIS Number will be reported as "FSISNOTSEQ".

WGS data must be interpreted within the context of how it will be used, e.g., to detect outbreaks or contamination events. Additional corroborating information, including case-patient food exposure and product distribution records, may be necessary to properly interpret the WGS data.

Prior Analysis

Prior analysis using this data can be found on the FSIS website, specifically the Ready-to-Eat (RTE) Meat and Poultry Sampling web page.

Information Available in NCBI's Pathogen Detection Isolates Browser

Single Nucleotide Polymorphism (SNP) cluster	A SNP cluster is a group of isolates whose genomes are closely related. In the Pathogen Detection Browser this element will contain a link that opens a new page with information about closely related sequences in the database.
Min-same or Min-diff	Minimum SNP distance from the query isolate to one of the same or a different isolation type. Isolation types are clinical (including human or animal) or environmental (including food).
BioSample	Further information (metadata) pertaining to the sample from which the sequence was isolated.
Assembly	Technical information pertaining to the sequence.
AMR Genotype	Information pertaining to antimicrobial resistant (AMR) genes found in the isolate sequence. Additional information about each AMR gene in this field is provided by a Reference Gene Catalog. Note: Empty cells do not necessarily indicate a lack of AMR genes.