FSIS Data Infrastructure Improvements: Public Health Information System (PHIS)

August 27, 2008

Carol Maczka, PhD
Assistant Administrator
Office of Food Defense and Emergency Response
Data Infrastructure Improvement

• Actively strengthening our public health data infrastructure to improve food safety and food defense.

• OIG December 2007 audit identified a number of areas for improvement in FSIS’ data infrastructure.
  – OIG has agreed to all 35 of FSIS’ responses to its recommendations.
Public Health Information System (PHIS)

- Improved information infrastructure will:
  - Integrate FSIS data streams
  - Support a data driven approach to FSIS inspection, auditing, and scheduling
  - Facilitate greater information sharing among external agencies
PHIS Modules

- Predictive Analytics
- Domestic Inspection
- Import
- Export
Predictive Analytics

• Will allow analysts to carry out ad hoc data analyses using multiple FSIS data sources in order to identify trends and anomalies
  • Example: Relationship between *Salmonella* test results and inspection findings

• Will monitor establishment data in real time and will have built-in alerts for anomalies
  Examples:
  • High rates of noncompliance for SRM in an establishment
  • Positive pathogen test result in an establishment (e.g., *E. coli* in ground beef, *Listeria monocytogenes* in ready-to-eat products)
  • Large number of inspection activities not completed in an establishment
Predictive Analytics

• Will have automated algorithms for consistent reporting and scheduling of inspection activities
  Examples:
  • Performance measure calculation
  • Risk-based sampling algorithms (*Salmonella*, *E. coli*, and *Listeria*)
  • Scheduling of Food Safety Assessments based on indicators of process control
  • District Activity Reports
  • Foreign establishment audit ranking

• Will have automated self-learning algorithms developed by Carnegie Mellon University that analyze data and create models to detect patterns in disparate data
  • Example: Analysis of relationship between pulsetype information from CDC’s PulseNet data on human illnesses and FSIS’ VetNet data on product sampling
Predictive Analytics Methods Development

• Question One
  – Can we use our methods, data, and tools to identify establishment risk factors that can be used to allocate resources based on public health impact?

• Question Two
  – Can we use our methods, data, and tools to discover new patterns in our data that might indicate problems?

• Question Three
  – Can we use our methods, data, and tools to support recall investigations and trace back?
Salmonella Data Sources

Positive samples

Serotype
PFGE Pattern
Antibiotic Resistance

Legend of Data Sources

FSIS
ARS
CDC
FDA

Regulatory Inspection data (PBIS, eADRS)
Microbial Sampling Data (M2K)
PulseNet Human PFGE Data
Est. info, and volume (PBIS, eADRS)
PHLIS Serotype Data
eLEXNET Food Safety Data
Question 1: Can we use our methods, data, and tools to identify establishment risk factors that can be used to allocate resources based on public health impact?

Analysis: Cross over FSIS/ARS Serotype data with CDC PHLIS serotype data to determine:

- Which serotypes in FSIS products cause the greatest amount of human illness.
- What percentage of *Salmonella* positives in FSIS regulated establishments are resistant to antibiotics.
The top serotypes of *Salmonella* causing human illness are not the same as those found most frequently in our establishments.
Analysis 2: What percent of *Salmonella* positives in FSIS regulated establishments are resistant to antibiotics.

This establishment has the highest count of positives

But this one has a much greater fraction of resistant isolates

Having the most positive samples may not make an establishment the riskiest to human health.
Question 1 Summary
We have identified refined risk measures that can be incorporated into the public health risk ranking algorithm:
-- serotypes of public health concern
-- resistance to antibiotics

A similar approach can be used to refine other variables of the public health risk ranking algorithm.
• **Question Two:** Can we use our methods, data, and tools to discover new patterns in our data that might indicate problems?
  – Analysis: Examined geographic and temporal trends in *Salmonella* antibiotic resistance using FSIS/ARS serotype data and CDC PHLIS serotype data.

• **Question Three:** Can we use our methods, data, and tools to support recall investigations and trace back?
  – Analysis: Combined FSIS/ARS data and CDC data to evaluate temporal and geographic relationships between *Salmonella* pulsetypes in FSIS products and those known to have caused human illness.
PHIS Component Status Update

- Currently in Design Phase
- System full production readiness Fall 2009
Domestic Inspection

- Automated capability to capture and report information on FSAs
- Ability to collect detailed information re: verifications, compliance, and other inspection-related activities
- Automated laboratory sample scheduling
- Secure data entry via internet
- Ability to pull information into our system from the inspection force, and from other internal and external sources
Imports

• More secure and timely transmission of foreign health certificates
  – Ability to receive electronic certificates from top three trading partners, provide advance notice and foreign government certification verification of US-destined product

• Integration with Customs & Border Protection’s ACE/ITDS

• Ability to schedule, track, and store results of foreign audits
Exports

• Automates manual processes (going from printed/handwritten forms to electronic)

• Automates edit checks to verify compliance with foreign import requirements

• Standardizes product descriptions and coding for products certified for export

• Provides ability for exporters to electronically pay fees