

Issue

National Advisory Committee on Meat and Poultry Inspection August 8-9, 2007

Data Collection and Analysis at FSIS: Standard Operating Procedures

Purpose

As discussed below, FSIS is developing a document to describe the current process for data collection and analysis (DCA) at FSIS. FSIS is seeking the Committee's comments on this document, as outlined in the questions below.

Background

FSIS' Data Analysis and Integration Group (DAIG) has developed a proposed process for data collection and analysis as described in the attachment. This document outlines the procedures for data collection and analysis, beginning with problem definition and including the technical plan, stakeholder input, and peer review. In addition, the roles and responsibilities of the groups that comprise the data collection and analysis teams (policy managers, Data Analysis and Integration Group, Data Coordination Committee, FSIS Program Offices) and how these groups interact also are described.

Questions

FSIS is presenting this draft outline to NACMPI to receive recommendations around the following questions:

1. Do you have any suggestions for improving our strategy for data collection and analysis?
2. Do you have other suggestions for stakeholder input in this process?
3. Do you have any other suggestions for conducting peer review?
4. Do you believe it would be worthwhile to form an on-going Sub-Committee to assist FSIS in evaluating various data issues? If so, please provide a rationale as to why it would be useful and recommendations on how it would be structured and should operate.

Contact Person:

Dr. Isabel Walls
Senior Scientist
Office of Food Defense and Emergency Response
Phone: 202-690-0834
Email: Isabel.walls@fsis.usda.gov

Data Collection and Analysis at FSIS: Standard Operating Procedures

**Food Safety and Inspection Service
Data Analysis and Integration Group**

Working Draft

July 30, 2007

Table of Contents

1.0	Introduction and Definitions	3
2.0	Roles and Responsibilities of the Data Collection and Analysis Teams	3
3.0	Process for Data Collection and Analysis.....	4
3.1	Problem Definition.....	4
3.2	Technical Plan.....	7
3.2.1	Summary of Issues/Questions.....	7
3.2.2	Identify, Collect and Review Data.....	7
3.2.3	Data Collection Strategy	8
3.2.4	Data Analysis Methods	9
3.2.5	Stakeholder Input and Peer Review	10
3.3	Collect Data, Perform Analyses and Develop Technical Reports	10
3.4	Use Data in Decision-Making.....	11
4.0	Program Evaluation	11

1.0 Introduction and Definitions

The Food Safety and Inspection Service (FSIS) is responsible for ensuring the safety and defense of the nation's meat, poultry and egg products, including domestically produced and imported products. Policy decisions for protecting the food supply need to be based on sound science. Valid and high quality data are the underpinnings of sound science. FSIS relies on a number of data sets to execute its food safety and food defense missions. The information contained in these data sets and any future data sets will be integrated within the new Public Health Data Infrastructure to provide support for the Agency's ongoing activities.

The process of data analysis at FSIS is evolving. The purpose of this document is to describe the current process for data collection and analysis (DCA) at FSIS. Specifically, this document outlines the procedures for data collection and analysis, beginning with problem definition and including the technical plan, stakeholder input, and peer review. In addition, the roles and responsibilities of the groups that comprise the data collection and analysis teams (policy managers, Data Analysis and Integration Group, Data Coordination Committee, FSIS Program Offices) and how these groups interact, are described.

Below are two key terms and definitions that have a bearing on this paper:

Public Health Data Infrastructure is the conceptual model for the FSIS technical environment; which includes the hardware (i.e. servers, PC's, laptops), system applications (i.e. AssuranceNet, Enterprise Reporting System, NRIMS, etc.), network, and communications.

Public Health System (PHS) is the major development project and system currently in the acquisition stage. The PHS will replace existing domestic (PBIS) and import (AIIS) systems with a new modeling and analysis modules as part of the core system.

2.0 Roles and Responsibilities of the Data Collection and Analysis Teams

To improve overall situational awareness and better inform decision making about food safety and food defense issues, FSIS formed a dedicated group, the Data Analysis and Integration Group (DAIG). The DAIG will characterize, coordinate, analyze and integrate data within and across different program areas. DAIG will: (1) ensure data analyses are relevant to Program Office business processes and the Agency's mission; (2) ensure data analyses are consistently of high quality; (3) conduct analyses to inform Agency decisions; (4) provide automated tools to facilitate data analysis and display; (5) conduct analyses to identify data gaps/needs within and across Program Offices; (6) develop sophisticated analytical models to integrate data streams and rapidly identify events, trends and anomalies; and (7) ensure data analyses are consistent with FSIS policies and OMB guidelines.

The DAIG works closely with a number of other groups and offices within FSIS. For example, the DAIG works with policy managers in FSIS to support decision making. The DAIG also relies on the input and expertise of the various Program Offices within FSIS. Consequently, a Data Coordinating Committee (DCC), comprised of senior Agency representatives, has been convened to serve as a liaison between the various Program Offices and the DAIG. Functions of the DCC include:

- Informing DAIG of the business processes of their respective Program Offices;
- Informing Program Office senior management of the activities of the DAIG; and
- Identifying subject matter experts to provide information for and feedback on data analyses conducted by DAIG.

The DAIG also works closely with the Office of the Chief Information Officer, administrators of the Public Health Data Infrastructure with emphasis on the Public Health System (PHS), to develop automated analytical tools to integrate data streams to rapidly identify events, trends and anomalies.

3.0 Process for Data Collection and Analysis

The goal of the data collection and analysis process is to provide the scientific basis for FSIS' decision making. A transparent process, based on sound science and inclusive of all stakeholder perspectives will improve the Agency's ability to effectively protect the food supply and public health. The overall DCA Process is presented in Figure 1.

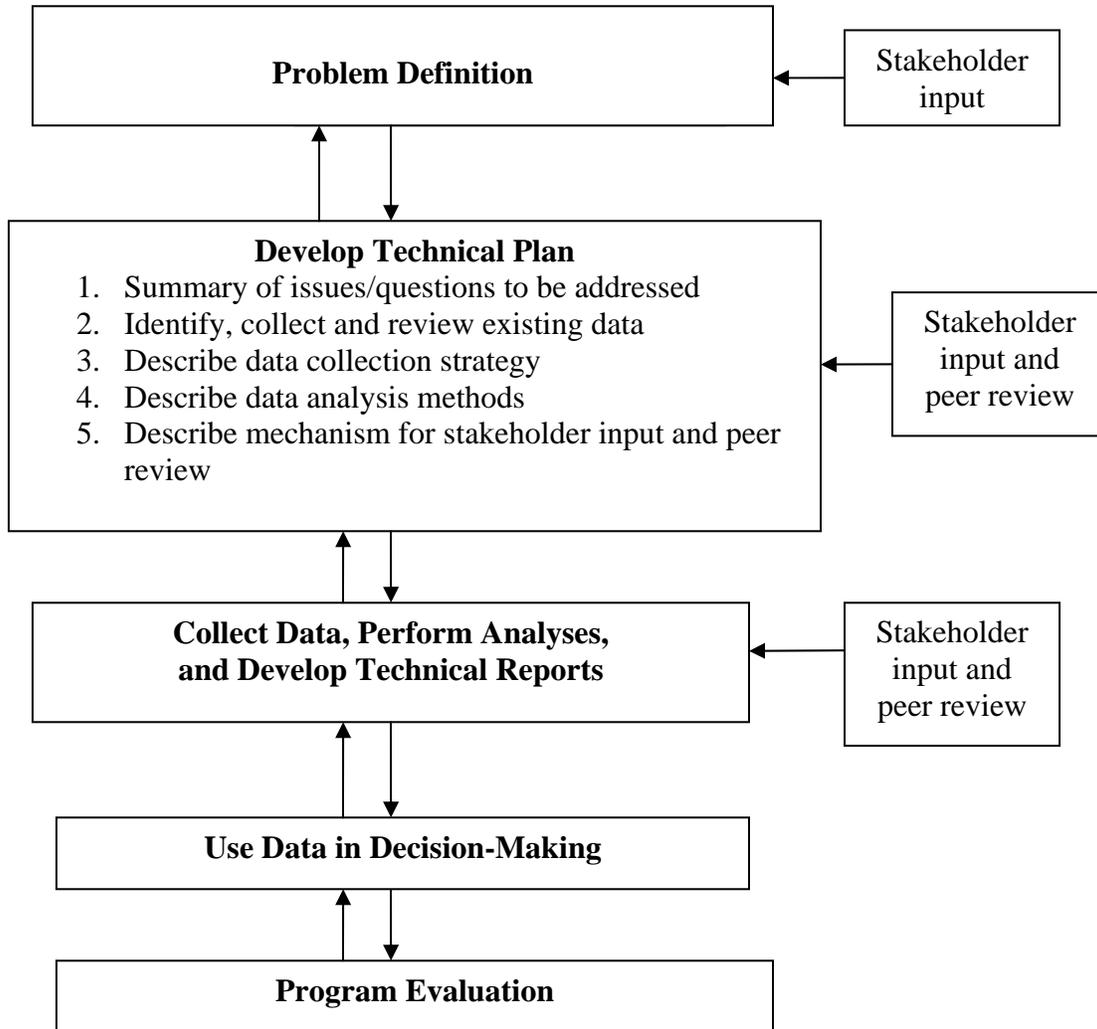
3.1 Problem Definition

The first step in the data collection and analysis process is problem definition. Clearly defining issues in terms of the questions to be answered helps to identify the criteria vital for decision making, increasing the strength and utility of the findings of the DCA process. Problem definition also helps to focus resources, minimizing the time and cost to perform the work required to develop recommendations.

Analytical questions typically are identified by policy managers in collaboration with data analysts and other Agency officials. Once the problems of interest are defined, the data collection and analysis options needed to inform decision making can be identified and prioritized. Variables include the kinds of information needed (e.g., quantitative, qualitative), sources of the information (e.g., documents, surveys, observations, interviews, focus groups), timing and frequency of data collection, resources available, and the timeframe of interest for decision making. Data collection should produce information that is correct, complete, valid, and unbiased. Problems that may be encountered in the process should be considered, along with mitigation plans. The data collection needs should be accompanied by a concise statement of the purpose and justification for each type of analysis proposed. Justifications should also consider the impacts on Agency resources.

Stakeholders will be offered the opportunity to provide comments during this first step in the process. Establishing a dialogue among data analysts, policy managers and stakeholders during problem definition will help to ensure that multiple perspectives are considered in the process. In addition, participation by all stakeholder groups will assist in providing the information needed to answer questions and reduce uncertainties.

Figure 1. Process for Data Collection and Analysis at FSIS



3.2 Technical Plan

The primary purpose of the technical plan is to describe the methods for collecting, reviewing and analyzing data to address the specific issues or questions identified during the problem definition phase. The choice of methods will depend on the questions to be addressed, the resources available, and the level of uncertainty acceptable for decision making. For example, information may be needed quickly if there is an outbreak of a foodborne illness and the Agency intends to recall the implicated products. Data collection activities also may be longer term; e.g., FSIS requires information on baseline levels of pathogens and needs to monitor how these levels change in products over time.

The technical plan for data collection and analysis will be developed by DAIG and presented to the policy managers. The policy managers and Agency officials will review the plan to ensure that it addresses the issues and questions identified during the problem definition phase. Policy managers may need to refine the issues or suggest changes in the proposed data collection and analysis objectives or timeline. Collaborative discussions between the DAIG analysts and the policy managers will be used to make any required changes to the plan. The plan will then serve as a guide to the analyst and a platform for obtaining stakeholder input and peer review. The Agency will invite public comment at this stage in the process and revisit the plan based on the input obtained. A call for data (based on data gaps identified) may also be made. External peer review of the plan will occur subsequent to stakeholder input.

In addition, the plan will specify the mechanism for stakeholder input and review, including internal reviews and external scientific peer reviews. Plan deliverables, a timeline for completion of each deliverable, and the team member responsibilities should also be stated.

The technical plan for data collection and analysis includes the following steps:

3.2.1 Summary of Issues/Questions

The technical plan will include a brief summary of the issues/questions to be addressed by the analysts, including descriptions of prior Agency actions relevant to the issue. Issues may include a variety of technical questions; e.g., risk-based inspection for processing establishments; risk-based sampling of products; pathogen reduction strategies; and program implementation questions received by the Technical Service Center.

3.2.2 Identify, Collect and Review Data

Once the policy questions have been developed, the next step is to identify and review any existing data from the PHS or elsewhere to determine how to best address the

issues/questions identified. All sources and types of existing data should be considered, including expert elicitation, empirical data, epidemiological data, and the results of predictive models. The data should be reviewed to determine whether they are representative and sufficient to answer the questions. It is important to recognize bias in existing data sets and to avoid bias when selecting data sets for use in decision making. A careful evaluation of the quality and limitations of existing data, including consideration of sampling design, methods used, analysis and presentation of results, must be done before accepting or rejecting information for analysis purposes.

In certain cases, it may be necessary to develop and document criteria used to determine the acceptability of available data sets. A weight-of-evidence approach should be used so that lower quality data sets are not given the same weight as higher quality data sets; e.g., data from peer-reviewed publications are preferable (and, therefore, more highly rated) than data from unpublished sources. Data collected more recently, or from the relevant geographic area, may carry more weight than older data or data from another region. References to any available data and other pertinent information should be listed. A review of existing data will provide information to determine whether and what type of new data are needed. In the absence of data, the Agency may elect to conduct a call for data, develop a strategy for acquiring more data, or conduct an expert elicitation to supplement existing data.

3.2.3 Data Collection Strategy

If new data are needed, they should be generated using standard methodology. A statistician should be consulted prior to collecting or analyzing data, to ensure that a statistically valid sampling plan is developed and followed. Validated analytical techniques should be used for microbiological and chemical testing. In particular, sensitivity and specificity must be known. If an expert elicitation is conducted in the absence of data, care should be taken to ensure that a balanced and unbiased panel of experts is used, representing all stakeholders, including consumer representatives and industry.

New data systems will be integrated with other Agency data in the PHS. The DAIG works closely with OCIO to ensure that new data streams can be integrated with the PHS as applicable so that the new data are reliable, secure, and accessible for analysis. The FSIS Data warehouse is the centralized repository for historical and statistical data, containing FSIS' data for decision support systems such as Enterprise Reporting System and AssuranceNet.

During the data collection process, the mechanism to be used for outcome evaluation will be considered, as the types of outcome evaluations that can be performed depend, in-part, on the types of data that are collected. Outcome evaluation will be planned at the same time as the data collection process, so that the data needed to support the evaluation can be included in the data collection plan.

3.2.4 Data Analysis Methods

The analytical tools and techniques used in data analysis will vary according to the issues/ questions being addressed. Knowledge of descriptive and inferential statistical techniques will enable extraction of useful information about populations based on survey samples. It will be important to describe variability and uncertainty of data. Probabilistic methods are useful for providing stochastic estimates, which may be preferable to deterministic point estimates, but they require information about the range of possible values for model parameters. However, stochastic models may not always be necessary; data analysts should clearly state the approach taken and the rationale. For example, assumptions may need to be made to save time, money, or resources, and these should be clearly stated to ensure that the uncertainties are recognized and the data are used in the appropriate context.

Some examples of the types of analyses that might be needed include:

Statistical Analysis - The statistical approach will depend on the nature of the data sets involved and the policy questions under consideration. This might include application of descriptive statistical methods, presenting summary statistics, statistical plots, curve fitting, trend forecasting, selecting or discarding certain subsets based on specific criteria, determining sample sizes, or other techniques. Both descriptive and inferential statistics and modeling may be used to arrive at conclusions.

Sensitivity Analysis - Sensitivity analysis will generally consist of varying model input parameters over a reasonable range (range of uncertainty in values of model parameters) and observing the relative change in model response.

Statistical Trend - Some policy questions concern time trends of sampled data. Statistical techniques can be used to perform time series analysis on selected data sets to identify time trends and statistically distinguish from random behavior.

Gap Analysis - Gap analysis can be used to determine if available data are sufficient for extracting useful information. Limitations of the data sets with respect to utility for policy issues of interest to the Agency will be identified.

Statistical Power - Statistical power is the probability of getting a statistically significant result given that there is a real effect in the study population. If particular a test is not statistically significant, it may be because there is no effect or because the study design makes it unlikely that a real effect would be detected. Power analysis can distinguish between these alternatives, and may be a critical component of formulating recommendations in response to certain policy questions.

3.2.5 Stakeholder Input and Peer Review

Stakeholder input and peer reviews are integral parts of the data collection and analysis process at FSIS and their use will be described in the technical plan. Public input will be provided by stakeholders (i.e., consumer groups, industry, academic research groups and other government agencies) during comment periods in response to Federal Register Notices and public meetings. It is important to obtain stakeholder input at different stages of the DCA process, since it may provide important insights into framing and context of the problem, information on additional data that may be available, and comment on the analyses and assumptions that underlie the Agency's plans to implement the policy recommendations. In addition, stakeholder input ensures that the Agency's decision making is transparent to all constituencies. Transparency is critical for credibility and scientific accountability. For example, data sources, data collection methods, variability and uncertainty in the data sets and decisions made based on the data must be clearly articulated. At the same time, the Agency must balance the need for protection of any sensitive or confidential FSIS-inspected establishment data that may be used in an analysis.

The Agency also believes it is important that independent scientific peer review be sought. There are a number of mechanisms which may be considered for external peer review and input including:

- National Advisory Committee on Meat and Poultry Inspection (NACMPI);
- National Advisory Committee on Microbiological Criteria for Foods (NACMCF);
- National Academy of Sciences (NAS); and
- Contractual arrangements with subject matter experts.

After the stakeholder input has been obtained and external peer review has been conducted, the concerns and comments raised will be evaluated by the appropriate staff. The Agency analysts will decide how to best incorporate the information obtained from stakeholder input and peer review in the DCA process. A document will be prepared that will address the disposition of each comment received and made available as part of the public record.

3.3 Collect Data, Perform Analyses and Develop Technical Reports

Data will be collected, quality-checked, stored and analyzed according to the methods described in the technical plan. Based on the results of the analyses, recommendations will be made on how to best address the policy issues/questions identified during the problem definition phase. Technical reports will identify the policy issues driving the analysis, the sources and quality of the data, methodology used, sources of uncertainty and variability, and data gaps and assumptions.

After completion, the draft analysis report will be presented to FSIS Program Offices as appropriate for comment, and revised accordingly. This should be followed by a public

announcement in the Federal Register and a meeting of the various stakeholders at which the findings of the data analysis report would be presented by appropriate personnel. The report should be revised in response to stakeholder comment and peer review.

The final report should effectively communicate the findings of the data collection and analysis process. The report should be written so that a technically trained third party could understand the results. An executive summary will communicate the intent, methods and results for managers.

3.4 Use Data in Decision-Making

Using the recommendations developed from the DCA process, FSIS will proceed with the development of Agency policy, Directives, Notices, regulations, operational procedures, or other actions that were the drivers for the policy issues/questions identified in the problem definition phase. Depending on the actions taken, a public meeting or other such outreach communication may need to be considered.

4.0 Program Evaluation

After a certain period of time following decision making, the Agency will collect information on the activities, characteristics, and outcomes of programs. This information will be analyzed and fed back to aid in refining program planning, development, and accountability.

The Office of Program Evaluation, Enforcement & Review (OPEER) conducts both formative and summative program evaluations. Formative evaluations focus on developing or improving programs. They are normally conducted during the development or ongoing implementation of new programs with the intent to improve them. Summative evaluations normally examine well-established programs in a much broader policy context.

Outcome evaluation has a role in both types of program evaluation work, and seeks to measure how well a program achieves its designed objectives. The stated goals of most (though not all) FSIS programs are expressed in terms of improvements in public health, such as reductions in foodborne illness. Given the difficulty of measuring changes in foodborne illness - especially attributable to a given type of food, Agency program, or establishment(s) - intermediate outcomes, such as changes in pathogen prevalence or changes in product recalls, are typically articulated and measured in lieu of direct public health outcomes.

Ideally, outcome evaluation should be conducted in a design framework that includes either a control or a comparison group. This is done to ensure that the presence - or absence - of desired program effects is not erroneously attributed to the program. For some programs, a “pre/post” design is the most feasible way to include a comparison

group. For others, a contemporaneous comparison group may exist in the form of entities (e.g., establishments) that are not included in the program being evaluated but are otherwise similar. In the design of the data collection effort- and the program itself consideration will be paid to how a comparison will be made and the types of data that will be needed to support the comparison.