Surveillance and Data Standards for USDA/APHIS/Veterinary Services

Version 1.0
July 11, 2006
Prepared by the Centers for Epidemiology and Animal Health

Table of Contents

Foreword

A primary objective of the National Animal Health Surveillance System (NAHSS) is to provide greater protection from endemic, emerging and foreign animal diseases that could affect the Nation's animal populations through enhanced information made available to decision-makers. Enhanced information is derived from accurate, valid and representative surveillance data. This document, *Surveillance and Data Standards for USDA/APHIS/Veterinary Services*, establishes a foundation from which to build surveillance and data management systems that will facilitate confident decisions.

The application of standards often generates varying levels of discomfort to those entities to which they are applied. However, the freedoms and flexibilities of current surveillance and data management systems have often hindered the ability to collate, validate and analyze surveillance data at the national level. The information presented in this document should be viewed as a roadmap toward achieving, through standardization, the accurate, valid and representative surveillance data required for a comprehensive and integrated NAHSS.

It is our intention that the standards presented in this document will initially be adopted by Veterinary Services and applied to facilitate the design of new surveillance and data management systems. We recognize that many of our current systems do not fully meet the standards and that concerted efforts will be necessary over time to transform the existing systems into function and format more useful for national surveillance. As work progresses on applying these standards, it is fully anticipated that revisions and additions to this set of standards will be ongoing to meet the dynamic needs of disease surveillance at the national level in the United States.

Introduction to Data Standards

Surveillance systems of the 21st century depend increasingly on the collection and storage of large quantities of data for quick, efficient access. However, collecting massive amounts of data is not synonymous with collecting useful and timely information, and is an inefficient and unreasonable substitute for well-planned, targeted surveillance. Planning surveillance systems in a standardized and methodical manner is essential to assure that the most appropriate information is collected and available to address the pertinent issues at the minimum cost.

Even when appropriate information is collected, its use can be complicated by storage in multiple databases, which may be administered by groups spanning Federal and State government, industry, and other non-government organizations. Further, data of poor quality and consistency in many cases are unreliable and may not provide useful information. For example, a database without standardization and quality control measures could report the data for a bovine disease with the species marked as “turtle” and the gender of breeding animals listed as “castrated.” Deriving conclusions and analysis from this information would obviously be difficult. Standardization will replace these errors with common, consistent business rules and coding.

Chapter 1 of this document provides standards and guidelines for the construction and operation of a surveillance system. These guidelines are intended to assist planners and managers in considering specific objectives, design strategies, reporting systems, implementation methods, and long-term system maintenance. The guidelines ensure that the objectives of the surveillance system are predefined, and that the collection, organization, and analysis of appropriate data is considered before implementation. Further, the guidelines allow for review and evaluation to ensure that the surveillance is providing the appropriate type and quality of information.

Standards for data categories and classes, in Chapter 2, provide guidelines for epidemiologists and database developers on the type and format of data to be gathered. These standards offer two major benefits. The first is the convenience of predefined classes that developers can use for data variables. For example, a pre-made list of breed and species codes may be quickly indexed and include suggestions for parameters such as variable lengths, types, and business rules. The second benefit is ease of communication between different databases to allow analysis of information from multiple sources.

Finally, standards for data storage and quality, in Chapter 3, ensure proper data entry and storage, and the proper structuring of data systems so they integrate readily with existing and future databases. It provides standards to guarantee data quality through validation and verification procedures, as well as training and entry guidelines. Further, the guidelines address accessibility to data users, while meeting requirements necessary for sensitive data. Chapter 3 also provides guidelines for data system documentation and changes to the existing system as indicated by changes in surveillance design, implementation, and technology.

Chapter 1 Standards For the Key Components of a Surveillance System

The information in Chapter 1 is intended to assist epidemiologists and surveillance experts who may be developing new surveillance systems or evaluating and improving existing systems. The guidelines provide an overview and generalized framework for details likely to be considered for comprehensive and efficient surveillance. The guide is not intended to be prescriptive or to define mandatory items for inclusion by surveillance architects, but instead to provide a useful tool to expedite development and review processes.

Four categories of information are included in the chapter (see Table 1). First is the introductory information that should lay the foundation for the surveillance system. These standards address components such as purpose, objective, and outcomes of the system, as well as basic considerations about the disease of interest and the people who collect, analyze, or base decisions on the resulting information. The second concept involves standards related to the design of the sampling strategy and data collection. Following this, a group of standards aid the surveillance professional in planning the appropriate analysis and reporting for the data collected. Finally, several guidelines are presented for the implementation and evaluation of the surveillance after a plan is developed.

Table 1. Summary of Standards For Key Components of a Surveillance System

Concept	**Standards:**
Introductory Information	1.1 Disease Description 1.2 Purpose and Rationale for Surveillance 1.3 Surveillance Objectives: Principal Uses of Data for Decision-Making 1.4 Expected Outcomes: Products, Decisions and Actions 1.5 Stakeholders and Responsible Parties
Population Description and Sampling Methods	1.6 Population Description and Characteristics 1.7 Case Definitions 1.8 Data Sources 1.9 Sampling Methods
Analysis, Reporting, and Presentation	1.10 Data Analysis and Interpretation 1.11 Data Presentation and Reporting
Implementation, Budget, and Evaluation	1.12 Surveillance System Implementation: Priorities, Timelines, and Internal Communications 1.13 Budget 1.14 Surveillance Plan Performance Metrics 1.15 Surveillance System Evaluation

1.1 Disease Description

Standard: The surveillance planning documents include current and relevant supporting information about the disease under surveillance.

a. Information included in the disease description is used to develop the case definitions. See standard 1.7.

Supporting Information:

The following classes should be included in the disease description section of the surveillance planning documents, or in documents describing the rationale of the surveillance system.

Class	Guidelines
Etiologic agent	➢ General categories: Virus, bacteria, toxin, external parasite, internal parasite, etc. ➢ Taxonomic classification: Family, genus, species; see <http://www.itis.usda.gov/>. ➢ Common name of the disease or condition. ➢ Pathogen strain or serotype.
Distribution	➢ Available information about location of current cases of disease should be identified. See location guidelines in Chapter 2. ➢ For FAD and diseases affecting trade, distribution throughout the world, at least by continent, should be included. ➢ For endemic diseases, distribution should be defined at the tertiary or lowest available administrative levels when possible, (e.g., National, State, or county.)
Clinical signs	➢ Description is targeted for the general practitioner audience.
Case definitions	➢ Criteria for a positive case, negative case, and others as applicable: Suspect, reactor, laboratory positive, etc. See standard 1.7, case definitions.
Pathological findings	➢ Necropsy results may be necessary for case finding, case definition, or passive reporting of cases. ➢ For lab-based surveillance and reporting, standards should be compatible with laboratory standards. See Chapter 2.
Epidemiology	➢ Considers and discusses the likelihood of disease introduction, outbreak, or change of status. ➢ Includes industry and management factors affecting disease transmission, (e.g., confinement operations, biosecurity practices, or industry awareness.) ➢ Provides measures or estimates of frequency, (e.g., prevalence, incidence, morbidity rate, mortality rate, or case fatality rate.) ➢ Transmission factors such as contagiousness, virulence of

	agent, or infectiousness, may be important components. ➢ Associated hosts, environmental conditions, and agent factors may influence the surveillance to be conducted. ➢ Susceptible species, population density, and location of the species are factors for conducting surveillance for many diseases.
Economic impact	➢ Economic impact compares the discounted long-term impact of not controlling the disease or conducting surveillance with the discounted long-term impact of controlling the disease. Includes the impact of disease eradication, assuming surveillance results in eradication, impact of government activities in the affected industry, and consumer and allied industry impacts resulting from the surveillance system. ➢ Economic indices of disease importance includes the direct and indirect costs of the surveillance system.
Methods for control	Mitigations and methods to control disease at national and herd levels. If conducting an economic impact assessment before initiating the surveillance system, the assessment should consider a number of scenarios about potential surveillance methods.

1.2 Purpose and Rationale for Surveillance

The purpose and rationale describe the need and reasoning for the surveillance system, and provide justification for the type of surveillance planned.

Standard: The purpose and rationale for surveillance are clearly described in the surveillance plan or equivalent program planning documents.

a. Responsible parties and stakeholders have a clear understanding of the purpose and rationale of the surveillance system.
b. The purpose and rationale are reiterated in progress and summary reports and presentations.
c. The purpose and rationale are periodically reviewed to determine relevance.
d. Changes in the purpose and rationale are documented and shared with stakeholders and all responsible parties.

Supporting Information:

The purpose describes the need and reasoning for the surveillance system and is justified by the rationale for conducting the surveillance.

Some of the purposes of surveillance are to:

- Estimate the magnitude and baseline status of a problem;
- Determine the geographic and demographic extent of an outbreak, predict possible spread, and provide data for disease regionalization;
- Describe the natural history of a pathogen or disease;
- Detect unusual clusters of disease (spatially or temporally), providing for early detection;

- Generate hypotheses and stimulate research;
- Define or assess the health status of a population, providing the foundation for market confidence;
- Detect changes in health practices, risk factors, or exposure;
- Facilitate planning of national control and eradication programs and strategies;
- Evaluate control measures and intervention efforts;
- Identify factors associated with a disease agent that may be used in conducting surveillance elsewhere and in modeling pathogen spread; and
- Determine times of year when most cases are observed.

The rationale for a surveillance system may include a description of the severity of the disease and its impact on trade, animal welfare, public health and other key areas. It may also include additional background information about the disease and its impact, and may summarize past surveillance efforts.

The purpose and rationale for the surveillance system may describe requirements for successful surveillance as well as measurement of success. Surveillance plans, supporting documents, and reports should use similar terminology to explicitly declare the purpose and rationale of the surveillance system.

The purpose and rationale of a surveillance system may change or evolve over time, and these changes must be documented and shared with stakeholders and responsible parties.

1.3 Surveillance Objectives: Principal Uses of Data for Decision-Making

Surveillance objectives identify the goals of the surveillance plan that will achieve the purposes described in 1.2 and outline how the resulting data and information will be used for policy actions or decision-making.

Standard: The surveillance objectives are specifically described in a surveillance plan or equivalent documents and explain the principal uses of the data for decision-making. They identify goals that when accomplished, will achieve the purposes of the surveillance system. Surveillance systems with multiple objectives identify and justify the relative priority of those objectives.

a. Responsible parties and stakeholders have a clear understanding of surveillance objectives and their relative priority.
b. Surveillance objectives are addressed in reports and presentations that describe progress in the surveillance system.
c. The relative priority of the objectives is demonstrated in the implementation of the system. Refer to standard 1.12.
d. Surveillance objectives and priorities are periodically reviewed to determine the extent of achievement.
e. Changes in surveillance objectives are documented and shared with responsible parties and stakeholders.

Supporting Information:
When surveillance objectives are achieved, the resulting data are or will be used for action. The following table presents the standard list of 10 surveillance objectives used as a

blueprint for developing swine surveillance systems. [1]

Foreign Animal Disease	➢ Rapid detection of FAD outbreak on domestic soil
	➢ Detect outbreaks of FAD on foreign soil
	➢ Monitor risk associated with domestic outbreak of FAD
Trade	➢ Document disease-free status
	➢ Describe disease prevalence patterns for regionalization
Disease control	➢ Assess progress in eradication and control campaigns
	➢ Assess progress in education campaigns
	➢ Assess progress in reduction of food-borne pathogens and zoonotic disease
Emerging Animal Disease	➢ Describe trends in hazards, exposures, and health conditions
	➢ Recognition of emerging diseases

Surveillance plans, supporting documents, and reports should use similar terminology to declare the objectives of the surveillance system.

The objectives of a surveillance system and use of its data for decision-making may change over time, and these changes must be documented and shared with stakeholders and responsible parties.

Priorities. Surveillance systems with multiple objectives must identify and justify the priority of those objectives. Criteria prioritization might include impacts on trade and productivity, animal welfare concerns, feasibility of control, cost of surveillance, and public health implications. Implementation of the surveillance system should demonstrate these priorities. See standard 1.12 for priorities in surveillance system implementation.

1.4 Expected Outcomes: Products, Decisions, and Actions

Expected outcomes include the information resulting from the surveillance effort, which is then used for decision-making, policy development, and action—as well as the physical products that are generated, such as databases and reports.

Standard: Surveillance planning documents articulate the expected outcomes of the surveillance system and describe the resulting products, including the decisions and actions resulting from implementation.

Expected outcomes:

a. Include specific statements for actions to be taken following identification of cases, and methods for ensuring timely response. This may be a reference to a VS response plan, to Uniform Methods and Rules (UM&R), or other actions that surveillance will trigger.
b. Include specific statements regarding actions to be taken when surveillance demonstrates freedom from the disease at the chosen threshold of detection, such as design prevalence or detection limit.

[1] Bush E, Lautner E, McKean J, Miller L. Swine Futures Project Final Report. U.S. Department of Agriculture Animal and Plant Health Inspection Service; 1999.

c. Include specific statements regarding how the surveillance system information will affect policy development and agency decision-making.
d. Include metrics allowing for evaluating performance of the surveillance. See standard 1.14 for surveillance plan performance metrics.
e. Are consistent with the objectives of the surveillance system.
f. Are a priority for those responsible for managing the surveillance system to ensure that the outcomes are realized.
g. Are periodically reviewed by stakeholders and responsible parties to compare these outcomes with the surveillance system products as part of internal system review. See document describing Surveillance System Evaluation Protocol; for more information, e-mail national.surveillance.unit@aphis.usda.gov.

Supporting Information:

The expected outcomes are statements that describe, using clear and specific terms, the tangible products that will result from the surveillance system. Expected outcomes include products such as data, reports, and information that will influence policy development or decision-making.

Examples of expected outcomes:

- "The surveillance system will produce an annual summary report describing detailed information on the status of DISEASE "X" in the United States…"
- "The surveillance system will provide timely and useful information to Federal and State decision-makers that will be used to determine what, if any, additional eradication efforts for DISEASE "X" are needed…"
- "The surveillance system will provide annual updates on the effectiveness of educational programs in reducing the incidence of DISEASE "X" in the States of a, b, c, and d…"

Statements of expected outcomes form the foundation for assessment of the surveillance system.

1.5 Stakeholders and Responsible Parties

Standard: The stakeholders and the individuals responsible for designing, implementing, managing, and disseminating information on the surveillance system are clearly identified in the surveillance planning documents.

a. Responsible parties and stakeholders are identified by role, title, group name, or agency rather than by name.
b. Responsible parties and stakeholders are consulted and engaged in surveillance planning and may, in some cases, contribute to the surveillance plan and procedures.
c. Responsible parties have appropriate and adequate training to fulfill their responsibilities. See Chapter 3.
d. An Information Technology (IT) team is identified early in the surveillance planning process; this team is responsible for developing and implementing the database system, including data entry and reporting requirements, for the surveillance program.

Supporting Information:

The documentation of responsible parties need not follow this format, but the individuals or groups responsible for these functions must be identified. The responsible parties should be engaged in the surveillance planning process to provide input regarding their expected role.

Class	Guidelines
Users of surveillance system information (stakeholders)	Are those who are: ➢ Policymakers ➢ Information users ➢ Beneficiaries of the surveillance information ➢ Data providers
Responsible parties for designing, implementing, managing, and disseminating information	Are the groups of individuals responsible for various aspects of the surveillance system, including: ➢ Surveillance system design ➢ Surveillance system implementation and leadership ➢ Data application design, development, and implementation ➢ Data application support and maintenance ➢ Data collection, including field data collection and laboratory data management ➢ Field data collection ➢ Laboratory testing ➢ Who detects cases ➢ Who confirms cases ➢ Who reports cases ➢ Field training of data collectors, data entry personnel, etc. See Chapter 3 ➢ Data analysis and interpretation ➢ Results dissemination and reporting ➢ Actions based on surveillance findings ➢ Review of surveillance system effectiveness

Stakeholders and responsible parties should be provided with appropriate information so they understand the expected outcomes of the surveillance system.

The IT team should be identified early in the surveillance planning process. This team identifies and articulates (1) business processes needed to capture inputs and produce appropriate outputs, (2) business rules for data collection, and (3) the risks, constraints, and assets of the data system.

1.6 Population Description and Characteristics

Standard: The population under surveillance is well defined in system planning documents and in system reports and publications. The population description inherently describes the scope or reach of the system (i.e., National, State, regional, local, and neighborhood.)

Supporting Information:

Depending on the structure of the system, at least two populations need to be described for most surveillance systems. The target population is the universe of eligible units at risk of the condition under surveillance, which gives rise to sampled units or cases. If the target population is sampled, then the study population should be described. The study population is the universe of sampled units that are investigated or counted in the surveillance system. If the target population is not sampled, then the catchment area of the population from which the reported cases come should be described.

The description of the population should include total size, animal type, administrative units, date(s) of surveillance, sampling design, and known risk factors. The following table provides more definition to these classes of a complete population description. See Chapter 2, data standards.

Class	Guidelines
Sampling units	➢ Should be identified and clearly defined. These may be simple units (individuals) or aggregates (herds or flocks). ➢ Geographic or spatial measures should be included, e.g., flocks per sq. km. ➢ Time constraints, if present, are included in description of the sampling unit.
Target population (not to be confused with *targeted* population)	➢ Population about which statistical inference will be made (general population at-risk) should be identified and clearly defined or estimated. ➢ If different from study population, the rationale for inference should be provided. ➢ Size of target population, e.g., number of herds by State, total number of animals by county, or population by Zip Code. (Note: Some NASS data are currently available at the Zip Code level.)
Study population	➢ The population from which the sample is to be drawn should be identified and clearly defined. In most cases, should be same as target population. ➢ Size of study population, (e.g., number of herds by State; total number of animals by county.) ➢ The sample frame (list of units to be sampled) from the study population should be identified.
Targeted population	➢ The population defined by specific disease variables inherent to the disease in question, (e.g., cattle with “high risk” clinical signs compatible with BSE.) ➢ The targeted population is intended to create intentional and predictable bias in the sample frame. ➢ If inference is made to the target (general) population of interest, a detailed explanation of the biological justification for the inference must be included.
Administrative units	➢ Define and include which units are included in the surveillance system , (e.g., States, regions, zones, counties, Zip Code areas, statistical reporting units, sample grid references, neighborhoods, and parcels.)

Size of sample service area	➢ Number of reporting units, (e.g., labs, clinics, slaughter plants.) ➢ Should include geographic area serviced per unit sampled. ➢ Number of eligible units served by reporting unit (per unit of geographic area being serviced).
Animal or group type	➢ The species, breed, and type (if applicable) of animals should be evident; include breeds and crosses. ➢ Define the animal by appropriate production phase concept. ➢ Age categories should include all appropriate categories pertinent to the surveillance objectives.
Date	➢ Population description should include applicable date. ➢ All date entries are as accurate as possible. At the minimum, include the date of sample collection and date of lab diagnosis. ➢ Report results are consistent using the defined date, (i.e., positive scrapie date is NVSL confirmed date.)
Sampling process	➢ Refer to standard 1.9, sampling methods.
Risk and exposure factors	➢ Population risk factors that may influence the outcome of the study. Confounders should be included in the description of the population (e.g., waste feeder hog operations). ➢ Risk factors need to be identified for statistical analysis. Anecdotal descriptions should not be used. ➢ When populations under consideration have unique risk factors or exposure to disease agents, the risk and exposure factors under study (independent variables) should be carefully defined to clearly differentiate the sectors of the population. If the surveillance applies to waste feeder, transitional, and commercial herds, risk/exposure factors must be clearly described prior to sampling.

1.7 Case Definition

Standard: The surveillance system has clear and understandable case definitions that include criteria for diagnosis, laboratory criteria for confirmation, any restriction or application of the case definition to specific geographical or demographic characteristics, and descriptions of case classification categories.

a. The working case definition is clear and understandable by the individuals who use it to identify and report cases.
b. Case definitions are consistent between all documents relating to a surveillance system or program, (e.g., CFR, UM&R, VS policy memos, etc.)
c. Case definitions are specific enough to avoid counting cases in more than one category.
d. Methods used to verify reported cases are clearly described.
e. The process is documented for handling data and information when case definitions change.

Supporting Information:

A surveillance system depends on clear case definitions for animal health-related events under surveillance. The case definition may include clinical manifestations (i.e., clinical signs); laboratory test results; necropsy findings; epidemiologic information such as subject, place, and

time; and/or specified behaviors, as well as levels of certainty including confirmed/definite, probable/presumptive, or possible/suspected).

The use of a standard case definition increases the specificity of reporting and improves the comparability of animal health-related events reported from different sources of data including geographic areas, and minimizes inappropriate regulatory actions, (i.e., movement restrictions, by placing cases in mutually exclusive categories.) Case definitions might exist for a variety of animal health-related events under surveillance, including diseases, injuries, adverse exposures, and risk factor or protective behaviors.

The following table provides guidelines on these classes of a case definition:

Class	Guidelines
Clinical description	➢ Brief one- to two-paragraph synopsis of clinical signs, history, and presentation. ➢ Acute, chronic, and late onset forms of disease should be described. ➢ Consideration should be given to asymptomatic or inapparent carrier reservoirs that could play a role in disease transmission.
Clinical case definition	➢ A clinical case definition can be used to broaden or restrict the sensitivity of a surveillance system by designating the species of animal(s) under surveillance and inclusion or exclusion of clinical signs or lesions for the disease or condition under investigation. Animals meeting a clinical case definition may be used to screen for inclusion of further testing.
Epidemiologic criteria and restrictions	➢ Criteria may restrict case definition to individual animals, herds, flocks, or premises that possess specific epidemiological characteristics. ➢ Criteria and restrictions may relate to the geographic location of an animal, farm, or premises; a particular point in time or season of the year; or a particular behavior associated with disease transmission or risk factor. ➢ Surveillance may be compartmentalized within a segment of a vertically integrated industry, (e.g., genetic grandparent stock, multipliers, commercial production); age group, (e.g., nursery vs. weaners); or commodity type, (e.g., meat type chickens vs. layers.) ➢ Criteria and restrictions should be used to clearly define population of interest under surveillance. See standard 1.6. ➢ May also include variables related to habitat, environmental conditions, seasonality, climate, etc.
Laboratory criteria	➢ May vary depending on the level of certainty required for surveillance. ➢ Screening tests are generally performed rapidly, are usually widely available within a laboratory system, and are relatively inexpensive. They typically trade lower specificity for higher sensitivity, which results in some level of false positive results. False negatives are undesirable but may occur. ➢ Confirmatory tests are typically less rapid and more difficult to perform, are less readily available within a laboratory system because of additional expertise needed to perform the test, and are more expensive than more commonly used screening tests. The ideal confirmatory test should be highly specific. ➢ Laboratory tests may not always serve as the gold standard for

	disease confirmation. Limitations should be identified and addressed. The type of diagnostic test and cutoff point or dilution used to define categories of cases may be included in the case definition. ➢ The type of test (e.g. ELISA, PCR, etc.), and any additional particulars specific to the testing, should be included if applicable.) ➢ The test should adapt to changing technology as new methods are established and are determined superior to older methods.
Case classification (definition categories)	➢ Case classifications usually represent various levels of certainty and include categories such as suspect, probable, and confirmed. ➢ Categories should be clearly defined and mutually exclusive. ➢ Levels of certainty may be defined using clinical signs identified through physical examination of the animal, antemortem or postmortem laboratory diagnostic tests, gross necropsy findings, histopathology, or the opinion of a recognized expert such as a foreign animal disease diagnostician (FADD). ➢ Reporting criteria may vary depending on level of certainty.
Required comments	➢ Surveillance plans may require certain forms of disease to be reported (e.g., fever, encephalitis, or meningitis); documentation of the vaccination status of the animal, herd, or flock; type of vaccine used; or lot number. Environmental exposure history of the animal, flock, or herd; history of arthropod exposure; or history of importation of animals, semen, or embryos from an endemic country or state may also be needed. Other information such as feedstuff or water source, evidence of foreign animal disease exposure or intentional release, or reproductive status could also be included in the case definition.

1.8 Data Sources

Standard: All data sources for the surveillance system are clearly identified.

Supporting Information:

Surveillance efforts typically rely on data from multiple sources. Some examples of data sources include:

- Producers
- Private practitioners or veterinary teaching hospitals
- Veterinary diagnostic laboratories, (e.g., university, State, Federal, or private)
- Government agencies, Federal, State, and local (e.g., animal health and agriculture agencies: FSIS, NASS, ERS, CSREES, FSA, RMA, RD, NRCS, etc., U.S. customs or border patrol, EPA, USGS, DHS, DHHS, Census Bureau, and others)
- Brand inspectors
- Affiliations or professional organizations, (e.g., milk marketing boards, registries)
- Businesses, (e.g., abattoirs, packing plants, pharmaceutical companies, zoos)
- Livestock markets
- Slaughter plants
- Renderers
- Business reporting services
- Surrogate/proxy data
- Aerial imagery

The surveillance documents should clearly describe all data sources that the surveillance system is expected to include, as well as associated strengths and limitations.

1.9 Sampling Methods

Standard: Sampling methods are thoroughly detailed in a surveillance plan or equivalent document.

a. The sampling methods include methods to assure: geographic representation without introducing spatial bias, an appropriate level of sampling, and a measurement method that ensures appropriate denominator information for analysis.
b. Sample size is determined with appropriate mathematical and epidemiological justification including measures of overall sensitivity such as design prevalence, detection limit, and level of confidence.
c. Sampling methods include information on modes of data collection, triggers for data collection, frequency of data collection, and transmittal of field or laboratory data to program managers or coordinators.
d. Sampling strategy should address and avoid unintentional bias, (e.g., sample selection, collection, reporting, and confounders.)
e. Test sensitivity, specificity, and predictive values are considered in the sample strategy.
f. Methods of data collection are described for each identified data source.
g. Sampling methods carefully consider data sensitivity and confidentiality issues. See Chapter 3 for data confidentiality standards. Applicable Federal regulations should be noted.
h. Changes to sampling methods for the surveillance system are appropriately documented and include a rationale for the change.

Supporting Information:

All sampling procedures and protocols, including field and laboratory data collection techniques, should be documented, including changes to data collection procedures as new procedures are identified to enhance the system. This information may be found in surveillance plans, uniform methods and rules, the code of Federal regulations, or annual or progress reports. See Chapter 3 for additional information on training staff for data collection and data confidentiality standards.

Population parameters are clearly described, including populations involved in the study and surveillance as well as populations for which inferences are being made. See standard 1.6, population characteristics.

Class	Guidelines
Sampling method	➢ Should be clearly defined and may utilize any justifiable epidemiological method, (e.g., simple random, systematic, cluster, stratified, or complex sampling, convenience or probability sampling.) ➢ Includes methods for randomization and stratification. ➢ Includes information on the use of grids and transects where applicable. ➢ Includes discussion, where applicable, of level of detection (threshold), statistical level of confidence, diagnostic (field) sensitivity of the sampling, predictive value, and in some cases comparison to other methods of sampling.
Outcome variables	➢ See standard 1.4, expected outcomes, products, decisions, and actions.
Exposure/risk factor variables	➢ See standard 1.6, population characteristics. Should be clearly identified.
Choice of sampling method	➢ Justification is described including purpose of the sampling strategy (e.g., sampling for detection, census, prevalence determination, or disease trends.)
Geographic extent of the study area under surveillance	➢ Sampling unit should contain a spatial context, (e.g., 2 randomly-selected samples per grid cell from 30 randomly-selected grid cells over a 400-sq.km. area, or 300 slaughter samples from abattoir that services 4 counties.)
Time intervals and frequency of data collection	➢ Sampling rate or frequency and the response rate for survey data. ➢ Time the agent, or the pathological consequences, are first observed. ➢ Information regarding date/times of sample collection, date/times of diagnostics, and date/times of lab reporting.
Methods of data collection and handling	➢ Refers to how raw data are gathered from the field, (e.g., face-to-face interviews, questionnaires, blood samples collected at sale or market, or necropsies of tissues from suspect animals.) ➢ Sample handling protocols, specimen chain-of-custody protocols, and cold chain measures should be available from laboratories. ➢ Sample degradation factors may be critical for some types of surveillance and should be addressed.
Sources of potential bias	➢ Are determined to all possible extent prior to data collection.
Trigger(s) for data collection	➢ Describe the event(s) that initiate data collection, such as the detection of an animal with clinical signs of disease.
Data collection and transmittal	➢ From the field or laboratory may include Web-based data entry forms, e-mail, fax, postal mail, spreadsheets (sent by e-mail or fax), or entry into database software such as Oracle or Microsoft Access.

1.10 Data Analysis and Interpretation

Standard: The methods used in summarizing, analyzing, and interpreting data are described in surveillance planning documents. In addition, the processes for analysis and interpretation are evident in reports and presentations.

a. The parties responsible for data summarization, analysis, and interpretation are identified.
b. A scientifically sound and detailed plan of data analysis and interpretation is consistent with the purpose, objectives, and expected products for the surveillance system.
c. Analytic methods are discussed in terms understood by VS professionals, including decision-makers whose expertise is not in a mathematical or analytical field. See standard 1.11.
d. The data type collected is appropriate for the method(s) of analysis planned (e.g., categorical data may not be amenable to some methods of quantitative analysis.)
e. Data analysis methodology is appropriate and supportable given the data sources, sampling methods, and type and quality of data.
f. Data interpretation provides timely, relevant information that meets surveillance objectives (standard 1.3) and expected outcomes (standard 1.4).
g. Methods for dealing with inherent biases, confounders, missing records, and unequal sample sizes are addressed.

Supporting Information:

Analysis: Outcomes of interest should have been determined prior to the collection of samples. General information is necessary for most analysis and may include descriptive statistics such as means, medians, modes, and standard deviations, as well as general epidemiological information including prevalence, incidence, and sampling duration. Of particular note to those developing surveillance plans and collecting data is consideration of population denominators. See standard 1.6, population characteristics.

More complex analysis, or analysis for unexpected purposes, may be conducted if the data are adequate; however, without advance planning, these needs may not be met.

Data interpretation is the process by which the analyst translates findings from the data into non-mathematical language useful for decision and policymakers. It should be transparent and describe the methods used to reach the options or conclusions presented. It also should be sensitive to the political environment, but the results not biased by political pressures. Where indicated, the interpretation of the analysis should provide options for decision-makers to consider. For example, if the analysis indicated a need for further surveillance, the analyst might provide parameters for different ways to achieve the goal along with strengths and weaknesses for each of them. Additionally, any assumptions that influenced the interpretation should be clearly articulated along with their ramifications and limitations. Finally, if possible, the sensitivity of any variable with exceptional influence should be discussed. For example, if the data from one area or set of samples have undue influence on the outcome, this influence should be discussed in the interpretation.

1.11 Data Presentation and Reports

Standard: Data presentation and reporting procedures specify the content, frequency, audience~~s~~, and methods of data dissemination (e.g., mail, e-mail, fax, private Web site, public Web site) for reporting data obtained through the surveillance system.

a. The parties responsible for data presentation and reporting are identified.
b. A plan for information dissemination is developed that assures communication to stakeholders and provides details on reporting format and frequency for distinct audiences so that communications are designed for maximal benefit for the target audiences.
c. Statistical or mathematic outcomes are explained in terms that an audience of non-specialists is likely to understand, and in technical detail adequate for peer reviewers.
d. Reports describing progress or conclusions are timely and relevant.
e. Presentations and reports carefully consider data sensitivity issues. See Chapter 3 for data sensitivity issues and references to Federal regulations regarding confidentiality.
f. The implementation of data reporting procedures is evident from reports and other presentations and publications.
g. Data reports and publications are consistent with the expected products for the surveillance system.

Supporting Information:

The dissemination of processed information derived from collected data to users and the linkage of targeted information to specific stakeholders completes the surveillance cycle. No surveillance system is complete unless it provides information that influences actions and decisions. Dissemination of surveillance data to those who need to know is a critical component of a surveillance system. Recipients should include those who prepare (or should prepare) reports, data collectors, and those with administrative or program planning and decision-making authority, as well as those involved with policy development.

A description of the surveillance plan and associated outcomes must include the intended audience and, for each audience, the communication format and frequency of reporting. Audiences include responsible parties within the surveillance system, as well as agency decision-makers and industry groups. In addition, the list of data sources should be considered as a specific audience. Careful planning is needed to target information and reports appropriately for maximal benefit.

The frequency of reports will depend upon the collection system used and the needs of users.

Reports should be timely and relevant. The frequency of reports will depend upon the collection system and the needs of users. Certainly, one must consider the impact and consequences of reporting data in varying stages of analysis and interpretation, but one must also weigh the impact and consequences of delaying reporting.

Chapter 3 provides additional information on data confidentiality standards.

1.12 Surveillance System Implementation: Priorities, Timelines and Internal Communications

Standard: The relative priorities of the surveillance system and timeline for implementing various aspects of the system are described in surveillance and implementation planning documents, and include specific information to facilitate internal communications.

a. The timeline for implementing the surveillance system is consistent with the stated priorities for the system and there is logical order in implementing the surveillance system.
b. Surveillance objectives have been prioritized and agreed upon by responsible parties.
c. Internal communication pathways and documents are clear and understood by all responsible parties.
d. All responsible persons or groups have received adequate training and have a clear understanding of their role in implementation.

Supporting Information:

The action steps needed to implement the various aspects of the surveillance system, and timelines, are described in the surveillance plan or associated implementation plan.

When surveillance objectives are met, information is provided to help determine actions and decisions. Some decisions and therefore some components of a surveillance system may have higher priority needs than others.

- The political environment and economic impacts associated with the disease may dictate the priority of implementation.
- Resource limitations and time constraints, as well as the need to pilot test various components of the system, may determine which components are addressed first.
- The objectives of the surveillance system should be carefully reviewed to prioritize implementation of various components of the system.
- To lend order to the prioritization process, stakeholders and responsible parties should agree on ranking criteria and their relative importance.
- The priorities of the agency, including those outlined in the VS Strategic Plan and NAHSS Strategic Plan, should be considered when prioritizing the surveillance objectives. See standard 1.3, surveillance objectives.

Responsible parties should develop appropriate documents for internal communications, including disease fact sheets, training manuals for all aspects of the surveillance system, and recording and reporting forms. In addition, internal communication plans and documents should be developed to ensure that all responsible parties understand the surveillance procedures and communication pathways, as well as the implementation action steps and timeline.

Communication with the IT application development team should occur early and often as surveillance moves from the planning stage to implementation. Responsible persons or groups charged with planning, implementation, and IT development should work cooperatively to develop clear data system requirements.

Training for data collection, data entry, sample collection, documentation, shipping of samples, and sample processing should be completed prior to commencement of each

individual's involvement in the surveillance activities. See Chapter 3.

1.13 Resources

Standard: Assessment of necessary resources is prepared to evaluate the human and financial aspects of design, implementation, and maintenance of the surveillance system.

a. The budget for the surveillance system is consistent with its priorities, purposes, objectives, and expected outcomes.
b. The budget for the surveillance system sufficiently ensures that the standards for data collection, management, and quality control may be achieved.
c. Human resources and technical expertise are available to achieve the surveillance priorities, purposes, objectives, and expected outcomes.
d. Budget information is routinely reviewed to evaluate alignment of the budget, the purpose and objectives of the system, and the system's products to date.

Supporting Information:

The budget includes direct costs.

Direct cost information includes any Federal funds appropriated to support other Federal or State agencies in the surveillance system. Information about use of budgeted funds – labor, rent, capital purchases, testing, indemnity, mail, supplies, vehicles, cleaning and disinfection costs, printing, etc. – must be collected.

Human resources are identified for the surveillance system. These may be new hires or redistributed from other duties.

1.14 Surveillance Plan Performance Metrics

Standard: One or more objective measures of the surveillance system's performance is included in the surveillance plan.

Supporting Information:

Metrics of performance should be part of a surveillance system and provide a means to measure the efficacy of the system.

- The performance metrics should be consistent with the objectives and expected outcomes for the surveillance system and, ideally, provide a measure of the extent to which expected outcomes are achieved.
- The metrics should be quantifiable and the unit of measure may be addressed for budgeting needs. The population and geographic scope of each metric is identified.
- The metrics may be modified or replaced over time as needed to meet the needs of the system or changes in technology.

Note: While quantitative performance metrics are ideal, some subjective metrics may also provide useful information for measuring a system's performance and these may be included as performance metrics.

Examples:

Testing adequate to meet specified prevalence at a given level of confidence. This might be a number of samples or a number of surveillance points from a targeted strategy.

Testing adequate to maintain a predetermined level of confidence. This might be equal to risk of disease introduction + the risk of pre-existing but not yet detectable disease.

Testing adequate to detect disease in a specified number of days/weeks/months.

Testing adequate to meet a specified sensitivity or specificity (i.e., probability of positive surveillance if disease exists or probability of negative surveillance if no disease exists).

Testing adequate to provide a negative predictive value (i.e., a measure of freedom from disease). This is the probability that no disease exists given negative surveillance.

A specified number of samples or observations in a given time, population, or location.

1.15 Surveillance System Evaluation

Surveillance system evaluation is the collection and review of information undertaken to assess how well the surveillance system fulfills its stated objectives and meets accepted standards. The evaluation process identifies system strengths and areas for improvement. The evaluation findings are intended to facilitate the system's role in a coordinated, integrated National Animal Health Surveillance System (NAHSS), consistent with the VS Strategic Plan.

Standard: The surveillance system is periodically evaluated to determine how effectively the system fulfills its stated objectives and meets surveillance system standards. The evaluation is conducted using the methods described in the Protocol for Evaluation of Surveillance Systems. (E-mail national.surveillance.unit@aphis.usda.gov. for more information.)

a. The purpose and outcome of the evaluation process are articulated and understood by responsible parties, stakeholders, and those involved in conducting the evaluation.
b. Results of the evaluation are shared with responsible parties and stakeholders.
c. Evaluation results provide meaningful information for program budgeting and setting program priorities.
d. Evaluation results provide viable alternatives, improvements, or suggested solutions for components of the surveillance system that are deficient.
e. The uses of past evaluation results are evident.

Supporting Information:

Surveillance systems should reflect national disease control and eradication priorities and promote the best use of public resources in the development of effective and efficient surveillance. National animal health surveillance systems have been developed in the past without standardized guidelines, designed at different times, and operated by different units within Veterinary Services (VS). VS field staff have been charged with implementing multiple systems with different objectives using different methods, terminology, and reporting forms and frequency. This may introduce extra costs and inefficiencies into these systems. Thus, there is a pressing need for regular evaluation of current animal health surveillance systems, especially given their complexity and the fact that most animal health data are collected under less than ideal circumstances. Also, in order to remain effective and useful, surveillance systems should adapt to changing situations such as new research findings regarding diagnostics, therapeutics or control procedures, significant changes in prevalence, legislation, global market pressures, or producer/public attitudes.

Surveillance systems should be evaluated regularly to ensure that they remain efficient, useful, and effective in order to meet their objectives. The evaluation process should include an assessment of whether the system's objectives are being achieved and whether it is serving a useful function. The evaluation of surveillance systems should include recommendations for improving quality and efficiency, such as eliminating unnecessary duplication.

The 2001 Animal Health Safeguarding Review recommended that VS promote a more coordinated and integrated approach to the surveillance and control of infectious disease. The proposed evaluation protocol is an important tool to achieve this goal.

Chapter 2 Data Concepts and Data Classes

2.1 Introduction

Animal health surveillance requires timely, accurate, and accessible data that facilitate analysis and reporting and contribute meaningful information for decision-making. Although surveillance databases are designed to collect specific data that will enable analysts to address the goals and objectives of a specific system, some data classes, or fields, are recorded in nearly all surveillance systems (e.g., dates, location information, species names, disease names, and others). The use of standards when capturing common data classes assures that the resulting databases not only provide data necessary to address system-specific objectives, but also for broader information inquiry and analysis. Within VS, data standards enhance our ability to provide timely and useful animal health information.

This chapter identifies the commonly used data classes in animal health surveillance systems and provides standard specifications for recording these fields. This chapter may also serve as a useful ”library” of common data classes to be considered when developing a data collection system for animal health surveillance purposes. Additionally, appendices at the end of this document provide specific codes for several commonly used data classes.

The data categories and classes described here do not include all fields needed to address the specific needs of a particular surveillance system, but provide a starting point for a data dictionary. Similarly, it is not expected that every data class listed will be useful for every surveillance system. Surveillance planners, data analysts, and database designers should determine their specific data needs and use the appropriate standards and specifications for their system. Business rules, entity relationships, and other ”best practices” in the design and implementation of database systems are beyond the scope of this chapter, although these factors may significantly impact data quality and accessibility. For surveillance systems within USDA/APHIS/VS, surveillance planners and database designers should request assistance from the VS Application and Information Management (VS AIM) group for more assistance with database design, implementation, and management.

Table: Summary of data concepts listed in this chapter.

Data Concept
2.3.01 Event
2.3.02 Subject
2.3.03 Population
2.3.04 Sample
2.3.05 Observation or Test Results Information
2.3.06 Premises
2.3.07 Person
2.3.08 Geographic Information

Terms. This chapter identifies *data concepts* and *data classes*. A *data concept* is a general grouping of data classes that are common to surveillance systems. *Data classes* are the individual data items (e.g., fields or variables) recorded within a concept. Although some concepts or classes are common to most surveillance systems, the use or importance of many will vary according to objectives and scope of the particular system. Data classes are distinguished into "tiers" based on frequency of use and how common they are in most surveillance programs.

The grouping of data classes into concepts facilitates the presentation of standards information so that data classes related to similar topics can be considered as a group. Data concepts may or may not parallel tables in a relational database. For example, classes in the concept "event" may all occur in a single database table, but classes in the concept "geographic information" will likely occur in many tables within a database.

Term	**Description**	**Examples**
Data Concept	A general grouping of classes for a surveillance system. These are generally non-redundant, non-ambiguous, and necessary for any surveillance system	➢ Event ➢ Sample ➢ Results ➢ Premises ➢ Person ➢ Geographic Info
Tier 1 Data classes	The most frequently used data classes in a concept	➢ Disease ➢ Species
Tier 2 Data classes	Data classes used in some but not all surveillance databases	➢ Breed

2.2 General Data Standards and Specifications

String values coded as integers:

- Some variables specified as string variables in this chapter may be coded as integers IF the database includes a related reference table to explain the code meanings (e.g., ”gender” may be coded as integers 1, 2, etc. when related to a reference table with a description for each code.)

- Boolean Values: For Boolean values that are coded as integers, no=0 and yes=1.

- Date fields: The standards format for date fields is mm/dd/yyyy

- Metric vs. English Measurement Units: Metric or English units may be used for measurement (e.g., weight), but the selected measurement system must be used consistently throughout the database.

2.3 Data Concepts

2.3.01 Data Concept: Event

Animal health surveillance data are typically collected in association with an event such as market testing, slaughter, on-farm testing, and others. Classes in this concept describe or are associated with animal health events.

Tier	Data class	Suggested Database Name	Description	Specifications (Type, length, precision, units)	Example
1	Event Type	EVENT_TYPE	Type of event (e.g., vaccination event, market testing)	String	Test; vac; insp
	Event Reason	EVENT_REASON	The trigger for the event or reason event is recorded	String or integer	Market Vaccination
	Event Date	EVENT_DATE	Date of the event	See “Date/Time” data concept	See “Date/Time” data concept
	Event Location (requires multiple fields)	EVENT_LOC	Location of the event; One or more fields that describe street address, city, county, state, latitude/longitude, or other location variables. The level of detail recorded for location will depend on management and analysis needs.	See “Geographic Information” data concept	See “Geographic Information” data concept
	Disease/condition	DISEASE	Disease or condition associated with the event; may be recorded as numeric code or string variable	String or integer	See Appendix A “Disease Codes”
	Species	SPECIES	Species associated with the event; may be recorded as numeric code or string variable	String or integer	See Appendix B “Species Codes”
2	Breed	BREED	Animal breed associated with the event; may be recorded as numeric code or string variable	String or integer	See Appendix C “Species-Breed Codes”

Initiating/previous event	---	Provides link to previous event	String or integer		
Investigator	---	Person who recorded the event information; often needed for program or data management purposes	See “Person” data concept	See “Person” data concept	
Person or Complainant	---	Person who requested or initiated the event	String		
Results Summary (may require multiple fields)	---	One or more fields that summarize the results from the event, (e.g., number tested, number negative, number positive)	String or double		

2.3.02 Data Concept: Subject

Data classes in this concept may describe an individual animal or an entire herd/flock. Surveillance analysis requires information about the characteristics of individual animals (or herds or flocks) and the population under surveillance. An individual animal record, or a record for an individual herd or flock, may include data fields that describe population characteristics; “subject” and ”population” fields may occur in the same database table. However, population characteristics are described in a separate data concept because those characteristics should be considered separately while planning data collection and database systems to ensure that the results will provide useful information for analysis.

Tier	Data class	Suggested Database Name	Description	Specifications (type, length, precision, units)	Example
1	Species	SPECIES	Species	String	See Appendix B “Species Codes”
	Individual or group indicator	---	Indicates whether the subject is a group or an individual animal	String or integer	
	Age	AGE	Numeric value of the subject’s age, usually in months or years	Single	5
	Age Unit	AGE_UNIT	Unit of measure for age, usually indicates month or year	String or integer	YR, MO 1, 2
	Animal Identification Type/s	ID_TYPE	Type of animal identification methods (e.g. tag type)	String or integer	See Appendix D “Identification Types”
	Animal Identification (may require multiple fields)	---	Variable(s) used to identify the animal or herd in the database (e.g., id numbers for the database record) and/or in the field (e.g., tag numbers)	String	---
2	Breed	BREED	Breed of the subject; may be recorded as a numeric code or string variable	String or integer	See Appendix C “Breed Codes”
	Gender	GENDER	Gender of the subject; may be recorded as a numeric code or string variable	String or integer	See Appendix E “Gender Codes”

Neutered	NEUTERED	Yes/ no / unknown field that indicates if the animal is neutered	String or integer (use 0=no and 1=yes when integers to code yes/no fields)	Yes/No/Unknown 0, 1
Age range (requires multiple fields)	---	When an age range is given for a herd rather than a single age value, the age range values should be reported in separate fields that indicate age and age unit values for the minimum and maximum of the range; use of text strings to describe age ranges should be avoided (e.g., avoid using “15 months to 5 years” to describe age range)	String and integer	Min_age=2 Min_age_unit=mo Max_age=10 Max_age_unit=yr
Subject location (may require multiple fields)	---	Subject’s (animal or herd/flock) current residence location One or more fields that describe street address, city, county, state, latitude / longitude, or other location variables. The level of detail recorded for subject location will depend on management and analysis needs.	See “Geographic Information” data concept	See “Geographic Information” data concept
Subject origin (may require multiple fields)	---	Subject’s (animal or herd/flock) location of origin One or more fields that describe street address, city, county, state, latitude / longitude, or other location variables. The level of detail recorded for subject origin will depend on management and analysis needs.	See “Geographic Information” data concept	See ”Geographic Information” data concept
Herd or Flock size	---	Herd or flock size is the number of animals in the herd or flock. This class is critical when the subject is the herd or flock rather than an individual animal, but herd/flock size may also be recorded when the subject is an individual animal or sample; see “population data” concept.	Integer	1422

	Weight	---	Numeric value of animal weight	Single	45.80
	Weight Unit	---	Unit of measurement for weight, often pounds, grams, kilograms; unit should be recorded uniformly throughout the surveillance database so that multiple words or abbreviations are not used to indicate the same weight unit (e.g., kilograms, kg, or kilos); also, use metric or English units consistently within a database	Integer	Lb, pounds, gm, kg, kilograms
	Use/Function	---	The intended use of the animal or herd from the owner's perspective	String	See Appendix F "Subject Use/ Function Codes"
	Status	---	Can be either an individual or herd/flock status	String	Vaccination, Weaning, Infected, Exposed
	Number Sick	---	When the subject is a herd or flock, numbers of sick animals should be reported by species on each premises	Integer	22
	Number Dead	---	When the subject is a herd or flock., number of dead animals should be reported by species on each premises	Integer	11
	Number Clinically Normal	---	When the subject is a herd or flock, number of clinically normal animals should be reported by species on each premises	Integer	33

2.3.03 Data Concept: Population

Population information is often captured in conjunction with event or subject records, and the specific population classes that need recording depend on the objectives and data analysis needs. Whether recorded explicitly during data collection associated with the surveillance activity or through other sources, details on size and demographics are essential for animal health surveillance analysis and should be considered.

Tier	Data class	Suggested Database Name	Description	Specifications (Type, length, precision, units)	Example
1	Species	SPECIES	Species of the population	String	See Appendix B "Species Codes"
	Herd/Flock Size	---	The number of animals in the herd or flock; size may provide important "denominator" values for analyzing surveillance data.	Integer	25,632
2	Age	AGE	Numeric value of age, usually in months or years	Single	5
	Age Unit	AGE_UNIT	Unit of measure for age, usually indicates month or year	String or integer	YR, MO 1, 2
	Breed	BREED	Breed of the subject; may be recorded as a numeric code or string variable	String or integer	See Appendix "Breed Codes"
	Gender	GENDER		String or integer	See Appendix "Gender Codes"
	Location (may require multiple fields)		Fields that describe the location of the study population	See "Geographic Information" data concept	See "Geographic Information" data concept

2.3.04 Data Concept: Sample

Data classes in this concept describe samples that are collected. Minimally, the sample identifier and sample type must be recorded.

Tier	Data class	Suggested Database Name	Description	Specifications (type, length, precision, units)	Example
1	Sample identification	---	Identifier used to uniquely identify individual samples; may be recorded as a numeric code or string variable	String or integer or double	A12345
	Sample type	---	Type of tissue sampled	String	See Appendix G “Sample Type Codes”
	Date collected	---	Date the sample was collected	See ‘”Date/Time” data concept	See ”Date/Time” data concept
	Collection site (may require multiple fields)	---	Location where the sample was collected; one or more fields that describe street address, city, county, state, latitude/longitude, or other location variables. The level of detail recorded for location will depend on management and analysis needs.	See ”Geographic Information” data concept	See ”Geographic Information’” data concept
2	Animal record identification	---	Key field used to link the animal and sample records in the database; required when animal and sample records are stored in separate tables	String or integer or double	
	Preservation	---	Method used to preserve sample	String	Ice, formalin
	Collector (may require multiple fields)	---	Person who collected the sample	See ”Person” data concept	See “Person” data concept

	Submitter (may require multiple fields)	---	Person who submitted the sample	See “Person” data concept	See “Person” data concept
	Method of shipment	---	Method of getting sample from collection site to testing site (lab)	String	Truck, air
	Date shipped	---	Date sample was shipped from collection site to testing site (lab)	See ‘”Date/Time” data concept	See “Date/Time” data concept

2.3.05 Data Concept: Observation or Test Results Information

Data classes in this concept describe results from tests performed on samples or observational data collected during an animal health event.

Tier	Data class	Suggested Database Name	Description	Specifications (type, length, precision, units)	Example
1	Test type	---	Type of test performed or test name	String	PCR, ELISA
	Test observation	---	Quantitative or qualitative values that describe the result of the test. The values recorded for test results, whether quantitative or qualitative, require careful consideration of potential broader applications of the data.	String or integer or single or double; depends on test type	Positive 1.2345 55
	Test date	---	Date test was conducted	See “Date/Time” data concept	See “Date/Time” data concept
	Test location (may require multiple fields)	---	Location where test was performed; may be a laboratory name, laboratory identification number, or multiple fields describing physical location. The level of detail recorded for location will depend on management and analysis needs.	String or integer, or See “Geographic Information” data concept	NVSL, or see “Geographic Information” data concept
2	Date received	---	Date sample was received	See “Date/Time” data concept	See “Date/Time” data concept
	Test date	---	Date the sample was tested	See “Date/Time” data concept	See “Date/Time” data concept
	Clinical Sign	---	Physical observation of symptoms. It is preferable to record individual signs in separate fields for electronic tabulation of data.	String	Recumbent, diarrhea, respiratory distress.

	Test interpretation	---	A qualitative judgment regarding the test outcome	String	See Appendix H “Result Interpretation Codes” or positive, suspect, negative
	System affected	---	Anatomical system affected by the disease or condition	String	Respiratory, reproductive, GI, CNS
	Person (may require multiple fields)	---	Person who interpreted or reported the test information; often needed for program or data management purposes	See “Person” data concept	See “Person” data concept
	Remarks	---	Remarks related to the test or observation	String	

2.3.06 Data Concept: Premises

Premises information may be captured in association with event or subject records, or captured separately and linked by common identifiers to event or subject data. Premises must be a physical geographic location (i.e,. post office box is not acceptable).

Tier	Data class	Suggested Database Name	Description	Specifications (type, length, precision, units)	Example
1	Premises id	PREM_ID	The unique identifier for the premises Premises identification numbers should be assigned using the National Animal Identification System	String Premises identification numbers should be assigned using the National Animal Identification System (alpha-numeric 7, random, character 7 is a check value)	002FVPL
	Location (requires multiple fields)	---	Location of the premises; this will include multiple fields describing physical address, city, county, state, latitude/longitude, and other location variables. The level of detail recorded for premises location will depend on management analysis needs	See “Geographic Information” data concept	See “Geographic Information” data concept
2	Premises type	---	A description of the type of premises, such as farm, slaughter plant, market, etc.		See Appendix I “Premises Type Codes”
	Species (may require multiple fields)	---	A description or list of species present on the premises; may be needed for program management purposes	String	See Appendix B “Species Codes”

Contact/s (may require multiple fields)		Premises owner(s) or contact person(s) for the premises; often needed for program or data management purposes; may include person name, phone numbers, mailing address	See “Person” data concept	See “Person” data concept	
Status	---	Status of the premises (e.g., quarantine, certified, exposed)	String	See Appendix J “Premises Status Codes”	
Active status	---	Current operational condition/status	String or integer or Boolean (yes/no)	Active Inactive	

2.3.07 Data Concept: Person

Classes in this concept describe the attributes of any person directly or indirectly related to the surveillance data or data flow such as data collectors, animal owners, laboratory personnel, and others. Classes in the person concept may occur in many tables related to the collection of animal health surveillance data. See the event, subject, observation or test results, and premises categories for instances where person is indicated as a data class.

Tier	Data class	Suggested Database Name	Description	Specifications (type, length, precision, units)	Example
1	Last Name	---	Person's last name; should be in a separate field from the first name	String, upper, no spaces	Smith
	First Name	---	Person's first name; should be in a separate field from the last name	String, upper	John or John Jacob
2	Agency Name (may require multiple fields)	---	Agency name or professional affiliation; names or acronyms/codes must be uniform throughout a database so that one agency is not indicated by multiple names or acronyms	String	APHIS
	Role / Concept/ Affiliation	---	The person's role or relationship to the regulatory program	String	
	Phone number(s)(may require multiple fields)	---	Area code plus 7-digit number	Number, xxx-xxx-xxxx	
	Fax Number	---	Area code plus 7-digit number	Number, xxx-xxx-xxxx	

	Address (requires multiple fields)	---	Person's mailing address	See "Geographic Information" data concept	See "Geographic Information" data concept
	Location (requires multiple fields)	---	Person's physical location; may be needed if physical location is different from mailing address	See "Geographic Information" data concept	See "Geographic Information" data concept
	E-mail address	---	Person's e-mail address	String; no spaces	person@agency.gov
	Person type	---	The person's role as it relates to the surveillance data or data flow	String	Owner, veterinarian
	Person Status	---	The person's activity status	String	Retired, active, inactive
	Prefix	---	Prefix to a person's name	String	Dr., Mrs.
	Suffix	---	Suffix to a person's name	String	Jr., Sr.

2.3.08 Data Concept: Geographic Information

Classes in this concept describe the attributes of any location information recorded in relation to surveillance data such as event or premises locations, or address information associated with people (see the "Person" data concept). Classes in the geographic information data concept may occur in many tables related to the collection of animal health surveillance data. See the event, subject, observation or test results, and premises categories for instances where location is indicated as a data class.

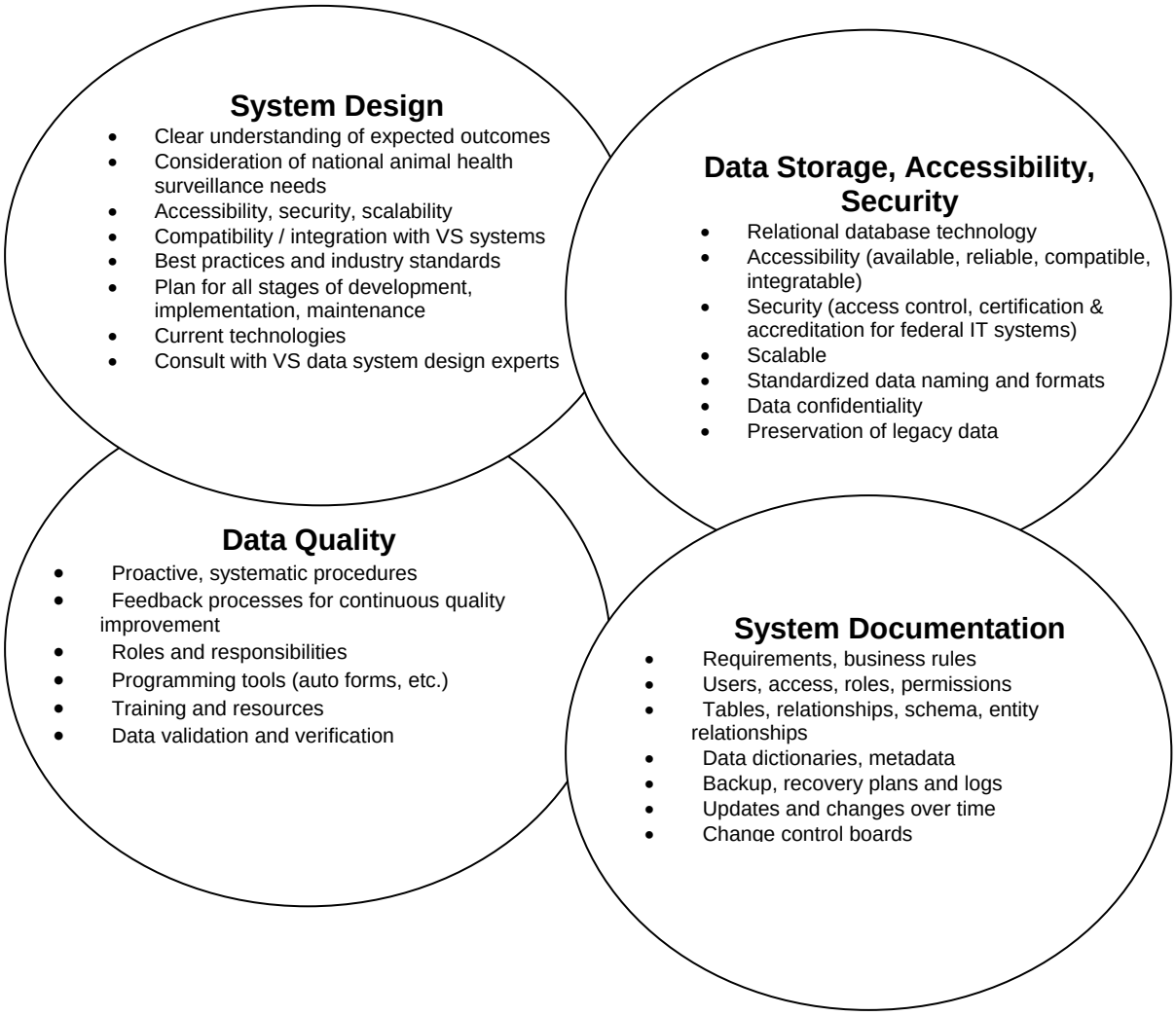
Tier	Data class	Suggested Database Name	Description	Specifications (type, length, precision, units)	Example
1	Address	ADDRESS	Physical address of the location	Use postal service standards	17454 Roller Coaster Rd
	City	CITY	City name	String	
	State	STATE	Two-character state abbreviation	Use postal service standards	AK, CO, NY
2	County		County name	String	Erie
	Zip Code		U.S. Postal Service 5-digit mail delivery area.	String	80132
	Zip Code plus 4		A geographic segment within the U.S. Postal Service 5-digit mail delivery area.	String	8312
	State / County FIPS code		State/County Federal Information Processing Standard Code	String	08041
	City FIPS code		Metropolitan Areas Federal Information Processing Standard Code	String	CO322
	Front Gate Latitude		Defined as the point of transition from public to private access when entering premises.	Double; decimal degrees	39.097886

Front Gate Longitude	Defined as the point of transition from public to private access when entering premises.	Double; decimal degrees	-104.845889
Coordinate type (front gate)	Method used to determine front gate coordinate (see appendix for values)	String	GPS, GEO See Appendix K "Coordinate Type Codes for Geographic Classes"
Validation code (front gate)	Method used to validate front gate coordinate (see appendix for values)	String	NVG, CBG See Appendix L "Validation Codes for Geographic classes"
Estimated positional error (front gate)	Estimated positional error during GPS data collection.	Double; meters	42

Chapter 3 Standards for VS Animal Health Surveillance Database Systems

The success of an animal health surveillance system depends largely on the availability of high quality and accessible data that can be easily and readily used for meaningful analysis and reporting. Data with these characteristics result from expert design of secure, accessible database systems; systematic procedures that assure data quality; and clear and efficient communication among a diverse group of professionals. That group includes, among others, IT experts, animal health surveillance experts, laboratory scientists, statisticians, program managers, veterinarians, and animal health technicians.

This chapter is intended to provide a high-level description of the standards for animal health surveillance database systems. While Chapter 2 focused on the details of standardizing database fields through the use of uniform naming and formatting conventions, this chapter addresses broader topics of database design, data storage, data security, data confidentiality, data quality, and related issues. The chapter describes expectations for designing, implementing, and managing VS data systems developed for national animal health surveillance. These expectations are organized around four broad and interconnected concepts: (1) system design, (2) storage, accessibility, and security, (3) data quality, and (4) system documentation.



The chapter does not provide detailed technical information on these subjects; users interested in technical aspects should refer to the wealth of books and Web sites available on these topics, or they may consult the professional experts within VS. The chapter also does not address data system needs for particular end users such as epidemiologists or geospatial analysts. The idea is that if the systems development concepts are properly addressed, then the needs of specific end users will be satisfied.

Table. Summary of Standards For Data Systems

Data System Design Standards:
3.1 Data System Design
3.2 Data Storage and Transfer
3.3 Data Security and Backup
3.4 Security Certification and Accreditation for Federal IT Systems
3.5 Data Quality
3.6 Data System Documentation
3.7 Data Confidentiality

3.1 Data System Design

Standard: The data system for the surveillance program ensures that appropriate data are collected and stored to meet the objectives of the surveillance program.

a. Integration and compatibility with other VS systems is carefully considered in the data system design.
b. The data system is designed using industry standards and best practices for database system development.

Supporting Information:

A sound database design creates the foundation of the surveillance system's data collection efforts. Database design ensures that the necessary data will be available to support the objectives and expected outcomes of the surveillance system and ensures the data will be available in a useful format. Resources describing best practices and industry standards for system and database design may be found in many books and Web sites, and the details of this subject are beyond the scope of this document. However, these considerations for VS animal health surveillance database design are emphasized:

- Data systems for VS animal health surveillance programs should be developed in consultation with VS' Center for Animal Disease Information and Analysis (CADIA). CADIA provides leadership in the areas of information systems, technology for collecting data in non-office environments, managerial and epidemiologic reporting, risk analysis methodologies, and computer-based modeling of animal disease.

- Data system design must consider integration and compatibility with other VS data systems including, but not limited to, Emergency Management Response System (EMRS), Generic Database (GDB has been recently renamed Animal Health and Surveillance Management information system [AHSM]), National Animal Identification System (NAIS), National Premises Identification System, and the VS Atlas of Animal Health Information - Spatial Data Library. When integration with existing systems is not possible, justification for this deficiency should be documented.

- Current technologies designed to enhance data communication and exchange between systems should be used to facilitate accessibility, integration, and compatibility (e.g., XML, HL7 messaging, web services). This concept relates to standards 3.1 and 3.2.

- Data system design should consider accessibility, security, scalability, maintenance/support, risks and impacts. The system documentation should reflect these issues (see standard 3.5, system documentation)

- As described in Chapter 1 (standards 1.10 and 1.11), a clear understanding of the objectives of the surveillance system, expected products of the surveillance system, and how data will be analyzed and reported should be established before data modeling is initiated. Application and database development teams should work closely with the surveillance planners and managers to assure that the data system is aligned with goals and expected outcomes.

- Data system design includes standard design practices such as:
    - Development and evaluation of conceptual and physical data models
    - Development of schema and entity-relationship diagrams that define relationships and reflect data flow, events, and business rules
    - Databases are normalized to reduce data redundancy and the associated pitfalls; ideally, databases are normalized to a modified third normal form.
    - Restricted use of user-defined or miscellaneous fields or tables

- Data system design follows standard stages of the database or software development life cycle, including planning, analysis, design, construction, testing, implementation, and maintenance.

3.2 Data Storage, Accessibility, and Transfer

Standard: Data for the surveillance program are stored in an electronic relational database using standardized naming and formatting conventions for data fields. The database system allows access to different types of authorized users and allows for easy data transfer and integration with other systems.

a. Electronic, relational databases are used to store data and are designed using the industry standards as described in standard 3.1.

b. To facilitate the flow of information among systems, the database uses the standardized data naming and formatting conventions described in Chapter 2.

c. The data are stored in a manner that facilitates the flow of information across systems.

d. The database should be scalable to allow for appropriate expansion of the system and reliable for continuous use.

e. The data storage system should maintain and make accessible legacy data for the program.

Supporting Information:

Relational databases. Electronic, relational databases provide the most accessible, reliable, scalable, and secure form for storing surveillance information. Spreadsheets are typically not appropriate for storing animal health surveillance system data. Relational databases provide these advantages:

- Tools for efficient and secure data storage, retrieval, and update
- Descriptions of the stored data items and their formats (i.e., metadata, data dictionaries)
- Tools for transaction and concurrency control that allow multiple users to access and use the database(s) simultaneously
- Tools for data recovery
- Tools for database security, including user authorization and backups.

Standardized data formatting. One of the challenges of database design and development is deciding what data to include (i.e., what fields or elements), and how to format the data

so it is compatible with other systems. Use of standardized naming and formatting of data fields greatly increases the ability to integrate information from multiple data collection systems. Chapter 2 addresses the naming and formatting of data elements.

Accessibility. Data and information accessibility is critical to a successful surveillance program. Animal health surveillance data must be accessible, in a timely manner, to a variety of authorized users who may need this information for analysis, reporting, and decision-making, particularly in emergencies. Tools should be available that allow access to data by authorized users so this information can be queried or extracted to other software for analysis (e.g., statistical or geospatial analysis).

Exchange of Data Among Systems. More information can be obtained from the same amount of data if the data are stored in a manner that facilitates the transfer of data to other systems or integration of data with other systems. Extensible Markup Language (XML) is commonly used for this purpose; XML facilitates information exchange among systems by encoding data with meaningful structure and semantics that computers and humans can understand (see <www.orafaq.com/glossary>). Other technologies or industry practices may also facilitate data flow among systems. This concept relates to standards 3.1 and 3.2.

Scalable. The database technology and architecture should allow appropriate levels of expansion of the system without compromising system efficiency.

Reliability. Reliability refers to the user's ability to consistently and efficiently use the database system for data entry or retrieval. Animal health surveillance systems must be able to consistently provide many users with simultaneous access to high quality data.

Legacy data (or historical data). Older data from some surveillance programs may have been collected and stored using outdated computer technologies. These data are often still a vital source of national animal health surveillance information and should be maintained for use by analysts and decision-makers. Legacy data should be accessible, and clear documentation should be available to describe these data to ensure usability in the future.

3.3 Data Security and Backup

Standard: Data for the surveillance program should be accessible to authorized users and be secure, recoverable, and reliable. The database is backed up to minimize intentional or accidental loss of information.

Supporting Information:

Security. The database should be strictly controlled and managed to prevent intentional or accidental loss of information. Security measures should include, but not be limited to, authorization and authentication of users, limiting privileges of users based on roles, encryption, data backups, and data storage in multiple locations. Security measures should address loss of data, loss of confidentiality, loss of privacy, loss of integrity, and loss of availability. The official process of data system security certification and accreditation for federal IT systems is described in standard 3.3.

Database backup. Backup is the process of periodically taking a copy of the database and moving it to an offline storage media[2]. Complete backup protocols and systems for animal health surveillance databases should be implemented to prevent loss of data that would effect daily operations. The backup system should consider how fast data recovery needs to occur, how long data must be available online, and methods for maintaining permanent backup media.

3.4 Security Certification and Accreditation (C&A)

Standard: All federal IT systems must be officially authorized, from a security standpoint, to operate.

Supporting Information:

Numerous directives explicitly state the requirement for owning agencies to perform a security certification and accreditation (C&A) of their IT systems. The C&A process involves a comprehensive evaluation of the technical, operational, and management controls in place to properly secure both the data and infrastructure of a given application. Accreditation is a formal acceptance, by the agency in question, of any residual risks once the security measures identified are in place, and is the official authority to operate the system as described in the associated documentation.

C&A standards require these steps:

- Establish and document access controls
- Maintain configuration management plan
- Perform a risk assessment
- Assess impact of data to personal privacy
- Ensure existence of a disaster recovery plan
- Create and maintain a system security plan

Critical components of the C& A process include:

- The system is certified and accredited every 3 years or following any major change to the system.
- The security classification of the system is properly assessed
- Required security documents are in place and updated as related attributes of the application evolve.
- The C&A process is integrated into the system development life cycle of any new system.

Information about the C&A process for federal IT systems is available through the VS Application and Information Management (VS-AIM) team; developers unfamiliar with the C&A requirements and process should contact VS-AIM.

[2] Connolly T, Begg D, Database Systems: A Practical Approach to Design, Implementation, and Management 3rd Edition. Addison-Wesley, 2002

3.5 Data Quality

Standard: The database system is designed, implemented, and maintained to ensure that data are of high quality.

a. Systematic and proactive procedures, used at multiple points in the data flow for the surveillance program, are implemented to ensure data quality; this includes feedback processes designed to not only inspect data for quality but also to address and fix the root causes of data errors.

b. Those responsible for data quality are identified and have a clear role in the data quality feedback process.

c. Programming tools are used to prevent data entry errors (e.g., code that enforces business rules or filters data that is out of range for particular fields).

d. Individuals responsible for data entry have training and resources to perform their job correctly.

e. Data are periodically validated and verified to ensure that the information is sufficiently accurate and complete to meet the needs of the surveillance program, and feedback from this process is used to improve the procedures used to ensure data quality.

Supporting Information:

Data that are complete and error-free are essential for a surveillance program to provide accurate reports and meaningful analyses in a timely manner. Data quality is defined as the ability of the data to meet the requirements of all internal and external customers. The level of accuracy, completeness, and timeliness may depend on the goals of a particular surveillance program, but animal health surveillance data collection efforts within VS also need to address the information needs described in the VS Strategic Plan [3]. In particular, VS' strategic goals for national animal health surveillance require data that are accurate, complete, accessible, available, and compatible with data from multiple systems. To achieve this, all animal health surveillance data systems within VS must demonstrate commitment to data quality. High quality data is not something that "just happens;" it is the result of systematic and proactive quality management efforts throughout the surveillance program and data system.

As with Data Systems Design (Standard 3.1), the measurement and improvement of data quality is a broad subject, and there are many books and Web sites on this topic. However, these considerations are of particular concern to data quality in VS animal health surveillance systems:

- Animal health surveillance systems within VS, and their underlying databases, should reflect commitment to implementing systematic and proactive procedures to

[3] "Veterinary Services Strategic Plan FY 2004 to FY 2008," http://www.aphis.usda.gov/vs/pdf_files/strat_plan.pdf. Accessed March 20, 2006.

assure data quality. These procedures should go beyond mere data inspection/correction and reflect a system-wide strategy to identify and eliminate causes of inadequate data quality. The allocation of resources within a surveillance program should reflect this commitment to data quality.

- Surveillance systems should identify teams of individuals who are tasked with data quality management; this group should include database managers as well as those involved with data entry, analysis, and reporting.

- Data entry user interfaces may be designed to minimize data entry errors. This may be achieved by restricting data entry in particular fields to acceptable values, cross-validating information within and between records, minimizing the use of free-text fields, using standardized terminology, using standardized lookup tables to control data entry, forcing users to double-key critical information, and constraints on data entry that assure data integrity.

- Training those responsible for data collection and data entry is recognized as a critical component to the data quality management process. The implementation portion of a surveillance plan should describe the process for initial and ongoing training of data collection and data entry personnel (Chapter 1, standard 1.12). Data entry personnel should be provided with resources (e.g., documents, protocols, user's guides, web-based resources, help desk contacts) to facilitate data entry and address questions or concerns. Feedback processes should be established to help data collectors and data entry staff identify and correct errors.

- Data validation and verification processes should be established to periodically review the data for accuracy, integrity, and completeness. These processes can identify systematic problems with a data system and avert data integrity problems that can affect quality reporting and analysis. These processes should not only identify and correct errors, but should also identify causes of data quality problems and provide solutions. Further, mechanisms should be in place to implement solutions to data quality problems when they are identified.

3.6 Data System Documentation

Standard: Data system documentation must be readily available and clearly describe:

- The system requirements and business rules
- The system users, including their access and roles
- The system components and their relationships (database schemas, entity-relationship diagrams, and descriptions of tables, views, and stored procedures)
- Data dictionaries that explain what data are stored in the database(s), how these are formatted
- Details regarding data backup and recovery procedures
- Changes to the data system over time, including changes to data elements that may impact analysis and reporting
- The activities of the system's Change Control Board (CCB), if applicable

Supporting Information:

Comprehensive system documentation provides critical information to IT developers, data end users, and other stakeholders about what, where, and how data are stored. System documentation must be developed for all animal health surveillance data systems. Examples of documentation for VS data systems can be obtained from VS CADIA.

Documentation of Changes to the Data System. Changes to a data system must be carefully considered and fully documented. Changes to any data system can have far-reaching ramifications both to developers of the system and stakeholders who use the data in the system, especially when data systems are integrated and pull or receive data from multiple sources. It is important to have a critical review process of all changes that occur to a data system's architecture and design. This includes, but is not limited to, changes in data collection, table design, table relationships, and coding systems used within tables.

Critical Components:

- A Change Control Board (CCB) should be established that has members representing the IT team and officials involved in planning and implementing the surveillance system. The CCB approves the original system, the system documentation, and system changes.
- Periodic reviews of the data system should be conducted to evaluate new technologies and modernization of the current system.
- New data systems within the agency should be evaluated for integration and contribution to the existing system.
- The data system should be evaluated periodically by unbiased experts to evaluate system efficiency and design.
- Any changes to the data system should be documented completely, including updates to the data dictionary, entity relationship diagrams (ERDs), and user guides.

3.7 Data Confidentiality

Standard: The data system ensures the confidentiality of sensitive or private information.

a. Appropriate authority is obtained to collect and maintain data that may be sensitive or confidential.

b. Sensitive or confidential data elements are identified and a protocol is developed and implemented to restrict access to these elements; access is limited to authorized individuals.

c. Staff are trained on data confidentiality policies and protocols.

d. Information-sharing policies for databases are documented and adhered to, ensuring confidentiality.

e. Federal and State information sharing laws and practices are carefully evaluated before linking, transferring, or sharing data with any organization.

Supporting Information:

Data confidentiality and privacy is a top priority for animal health surveillance systems. Surveillance systems must include approved protocols for collecting, storing, restricting access to, and working with confidential information. In addition, confidential information may only be shared when the situation complies with Federal and State laws.

First steps for addressing data confidentiality issues include:

- Obtain authority to gather and maintain private information
- Identify data that require restricted access
- Identify groups of users (roles) and the data they can access
- Restrict user access to data based on group membership (roles)
- Determine if information on individuals or groups of individuals can easily be obtained
- Ensure the information stored is used as authorized

For VS data systems, refer questions on data confidentiality and information sharing to APHIS Legislative and Public Affairs.

Glossary of terms for surveillance standards

Case Definition: A set of diagnostic criteria that must be fulfilled in order to identify an individual as a case of a particular disease. Case definition can be based on clinical, laboratory, or combined clinical and laboratory criteria, or a scoring system with points for each criterion that matches the features of the disease.

Certification and Accreditation Standards: (Security C & A Standards): Process involving a comprehensive evaluation of the technical, operational, and management controls in place to properly secure both the data and infrastructure of a given application. Security and accreditation is required for all federal information technology systems.

Change Control Board: Consists of database developers and surveillance officials who are responsible for the analysis, approval, rejection, scheduling and control of the implementation of any changes to a surveillance data system.

Clinical Case Definition: A clinical case definition can be used to broaden or restrict the sensitivity of a surveillance system by designating the species of animal(s) under surveillance and inclusion or exclusion of clinical signs or lesions for the disease or condition under investigation. Clinical case definition may be used to screen animals for additional testing.

Data Accessibility: In this document, data accessibility refers to the level at which data may be easily accessed and used by the individuals who need the information for surveillance purposes.

Data Classes: The individual data items recorded within a concept.

Data Concepts: Separate divisions designed to group together data that share common attributes.

Data Confidentiality: The practice of protecting sensitive data from unauthorized disclosure.

Data Quality Control: The process of ensuring that data recorded for a surveillance system are free from errors. The data quality control process includes data inspection, validation, verification, and review of data for accuracy, integrity, and completeness.

Data System: For the purposes of this document, a data system refers to all elements used for the management and storage of the data associated with a surveillance system. This includes software and hardware; relational databases; and tools used for database administration, data management, data entry, reporting, exchange of data with other systems, and analysis.

Epidemiology: The study of a disease pattern in a population to determine prevention and control strategies.

Event: The places associated with the collection of animal health surveillance data, including market testing, slaughter, on-farm testing, and others.

Expected Outcomes of Surveillance: The expected results of a surveillance system. Expected outcomes include information to be used for decision-making, policy development, and action as well as the physical products anticipated to result from the surveillance effort such as data, databases, reports, and information.

Geographic Information: Describes the attributes of any location information recorded in relation to surveillance data such as event or premises locations or address information associated with people.

Observation: Something perceived by the senses of an individual.

Person: Any person directly or indirectly related to the surveillance data or data flow such as data collectors, animal owners, laboratory personnel, and others.

Population: The whole collection of units from which a sample may be drawn; not necessarily a population of individuals--the units may be institutions, records, or events.

Premises: The physical site where the subject(s) of an investigation or sample collection are located.

Responsible Parties: The groups of individuals responsible for various aspects of the surveillance system, including those responsible for: design, implementation, leadership, database development, management, support and maintenance, data collection, training, analysis and interpretation, dissemination and reporting, actions based on surveillance results; and those who review and evaluate the surveillance system.

Sample: A specimen of fluid, blood, or tissue collected for analysis on the assumption that it represents the composition of the whole.

Sensitive Data Elements: Records that need to be protected according to State and Federal Privacy Act laws.

Standardization: A generalized framework for details most likely needing to be considered for comprehensive and efficient surveillance.

Subject: An animal subject to treatment, testing, observation, or experiment.

Surveillance: The activities involved in the systematic collection, collation and analysis of animal health data combined with the prompt dissemination of vital information to those who might take action.

Surveillance Plan: A documented framework that systematically describes the disease, including the purpose, rationale, objectives, and outcomes; stakeholders and responsible parties; population and sampling methods; performance metrics; plans for analysis, reporting, and presentation; and expected implementation, budgeting, and evaluation plans needed to develop a surveillance system.

Surveillance Stakeholders: Users of surveillance system information**.**

Surveillance System: A comprehensive and coordinated system that will collect, collate, and analyze animal health data and promptly disseminate animal health information.

Surveillance System Evaluation: The systematic collection and review of information about a surveillance system undertaken to assess how well the system fulfills its stated objectives and meets accepted surveillance standards.

System Documentation: Detailed records of database development and maintenance that allow long-term data quality and integrity. This documentation is important for relaying system components to stakeholders, users, and those maintaining the system. Quality documentation helps minimize problems associated with changing staff and changing system requirements.

System Security: Processes undertaken to prevent the accidental or intentional loss of information, or the unauthorized access to or use of information.

Validation of Data: The specific process of checking data records to detect specific errors; validation involves reviewing data against a set of criteria used to determine if values are logical or within acceptable ranges.

Verification: A systematic process of determining whether collected data are sufficiently complete and correct to meet the requirements for the surveillance system.