GOODSamples: Guidance On Obtaining Defensible Samples

© Bruce Stambaugh, 2012.

GOOD Samples

Sampling and Sample Handling Working Group
FDA, AAFCO, AFDO, APHL, and Industry

October 2015
http://www.aafco.org/Publications/GOODSamples

© 2015 Association of American Feed Control Officials (AAFCO)
1800 S. Oak Street, Suite 100, Champaign, IL 61820-6974
http://www.aafco.org/

Cover photograph used with permission of Bruce Stambaugh (brucestambaugh.com). Taken at Mt. Hope Produce Market, Mt. Hope, Ohio.

Table of Contents

Continued

PREFACE

The development of this guidance document, *GOODSamples*, was a collaborative effort by representatives of the US Food and Drug Administration (FDA), food and feed state regulatory agencies, and industry. The impetus for the effort is to improve analytical data equivalency among state, federal, and local agencies to enable inter-agency sharing of data collected in support of food and feed regulatory programs. Because analytical data is only as good as the the quality of the sample, it is essential that any improvements in the measurement process begin with the selection of the primary sample. In addition to improving equivalency, the guidelines provide the foundation for defensible decisions.

This work was developed by a working group established under Cooperative Agreement #U18FD004710-01 and composed of members of the Association of Public Health Laboratories (APHL), Association of Food and Drug Officials (AFDO), the Association of American Feed Control Officials (AAFCO), and industry. This project was 100% funded with federal funds from a federal program of $1.3M.

APHL

Association of Food & Drug Officials
AFDO
Founded 1896

Food and Feed Industries

AAFCO
Association of American Feed Control Officials

CONTRIBUTORS

Working group chair: Nancy Thiex, consultant to AAFCO, Thiex Laboratory Solutions LLC, Brookings, SD

Subject matter expert: Charles Ramsey, EnviroStat Inc., Windsor, CO

Working group members:

Rina Bhikha, FDA Office of Regulatory Affairs, Long Beach, CA

Jo Marie Cook, Florida Department of Agriculture and Consumer Services, Tallahassee, FL

Andy Crawford, consultant to AAFCO, Crawford Science Consulting, Hacienda Heights, CA

Dan Danielson, Tennessee Department of Agriculture, Nashville, TN

Qian Graves, FDA Center for Food Safety and Applied Nutrition, College Park, MD

William Hirt, ANSI-ASQ National Accreditation Board, Alexandria, VA

Gregory Hoffman, Zoetis, Kalamazoo, MI

Lourdes Andujar, FDA Office of Regulatory Affairs, Aguada, PR

Michael McLaughlin, FDA Office of Regulatory Affairs, Rockville, MD

Kenneth McManus, Maryland Department of Agriculture, Annapolis, MD

Sarah McMullen, FDA Office of Regulatory Affairs, FDA Southeast Regional Laboratory, Atlanta, GA

Angela Montalbano, New York State Department of Agriculture and Markets, New York, NY

Lawrence Novotny, retired from South Dakota State University, Brookings, SD

Yvonne Salfinger, consultant to AFDO and APHL, Denver, CO

Barb Schroder, Minnesota Department of Agriculture, St. Paul, MN

Thomas Scott, FDA Office of Regulatory Affairs, Winchester Engineering and Analytical Center, Winchester, MA

Richard Stephens, Florida Department of Agriculture and Consumer Services, Tallahassee, FL

Cooperative agreement liaison: Donald Burr, FDA Office of Regulatory Affairs, Rockville, MD

Technical reviewer: Francis Pitard, Francis Pitard Sampling Consultants LLC, Broomfield, CO

Technical editor: Louise Adam, FASS Inc., Champaign, IL

INTRODUCTION

GOODSamples outlines the scientific and systematic approach to ensure that analytical data generated as a result of a sampling process is representative of the decision unit and is defensible. The intended audience is anyone involved in food and feed safety; however, the principles are generally applicable to any similar industry. Considerations for sampling procedures include the determination of the sample quality criteria, determination of material properties and application of sampling theory. The resulting sampling protocol will specify minimum mass/volume needed, minimum number of increments, selection of increment location, sample integrity requirements, sampling tools and equipment, and quality control, resulting in defensible measurement data. *GOODSamples* does not provide specific sampling protocols. *GOODSamples* does provide the scientific and systematic approach needed to develop or evaluate sampling protocols for defensible decisions. Following the process outlined in *GOODSamples* will result in increased confidence in the safety of food and feed.

Definitions and Acronyms

The concepts presented in *GOODSamples* for guidelines in support of a systematic, harmonized approach to sampling cannot be communicated without a workable vocabulary. Terminology in the sectors involved in sampling is very diverse; thus, it becomes necessary to establish critical terms with definitions to ensure that readers are able to comprehend the message as intended by the authors. The following terms have been determined as the minimum number of terms to effectively communicate the concepts herein. Previous notions for these terms may need to be abandoned in favor of the definitions provided here for effective understanding. It may be helpful to keep this list handy as you study *GOODSamples*.

Analyte integrity: The characteristic or concentration of the analyte of interest is maintained from collection of the primary sample through the test portion (maintain sample correctness).

Bias: The tendency for a measurement to systematically over- or underestimate the actual (true) value.

Comminution: Reduction of particle size by crushing, chopping, blending, and grinding, among others.

Compositional heterogeneity (CH): The heterogeneity arising from differing composition among individual elements (e.g., particles) in a decision unit.

Decision unit: The material from which a sample is collected and to which an inference is made.

Distributional heterogeneity (DH): The heterogeneity arising from the non-random spatial or temporal distribution of elements within a decision unit.

Element: The individual components that comprise a material (e.g., particles or fragments for solid materials, molecules for liquids, particles and molecules for slurries).

Finite element materials: Materials composed of elements that can be individually identified and individually selected at random.

Infinite element materials: Materials composed of elements that cannot be individually identified nor individually selected at random.

Evidentiary integrity: Demonstration that the analytical result(s) can be traced to the decision unit and have not been compromised. In legal terms, it is the identification and authentication of the evidence.

Global estimation error (GEE): Total errors in the entire measurement process, from primary sampling through final measurement.

Imprecision: The tendency for a measurement to vary randomly from the actual (true) value. Imprecision describes data spread or dispersion. There are two major sources of imprecision error: fundamental sampling error (FSE) and grouping and segregation error (GSE).

Increment: An individual portion of material collected by a single operation of a sampling device and combined with other increments to form a sample.

Inference: The process of estimating a concentration or characteristic about a larger amount of material from data derived from a smaller amount of material.

Lot: A specific identified portion of a batch, having uniform character and quality within specified limits; or, in the case of a product produced by continuous process, it is a specific identified amount produced in a unit of time or quantity in a manner that assures its having uniform character and quality within specified limits (21CFR210.3).

Mass reduction: The process of selecting a smaller mass from a larger mass (not to be confused with comminution/particle size reduction).

Nugget: An element that is largely different from neighboring elements and causes extreme compositional heterogeneity.

Probabilistic sampling: A sampling strategy in which all material has an equiprobable chance of being collected.

Sample: A portion of a material selected from a larger quantity of material. The word "sample" should only be used with a modifier as follows:

Primary sample: The collection of one or more increments taken from a decision unit according to a sampling protocol.

Laboratory sample: The material received by the laboratory.

Analytical sample: Results from any manipulation of a laboratory sample.

Test portion: The quantity of material taken for measurement.

Sample *(continued)*:

Replicate sample(s): Additional samples collected under comparable conditions at any point in the sampling process.

Split sample(s): Portions obtained when a primary, laboratory, or analytical sample is divided into equal portions.

Composite sample: A mixture of primary samples or laboratory samples, combined before analysis for the purpose of analytical efficiency. This term is often misused and is, therefore, avoided in this document.

Sample correctness: A condition achieved when bias is controlled to a negligible level. Major sources of bias include increment delimitation error (IDE), increment extraction error (IEE), and increment weighting error (IWE). Sample correctness is a necessary condition for a representative sample.

Sampling protocol: A sampling protocol is a detailed procedure for obtaining a representative sample from a specific decision unit that meets the sample quality criteria. The protocol includes appropriate mass, number of increments, sample correctness, quality control, and procedures for maintaining evidentiary integrity.

Sample quality criteria (SQC): A series of statements that clarify program technical and quality requirements to support defensible decisions. These statements include the question to be answered, definition of the decision unit and the desired confidence in the inference.

Theory of sampling (TOS): Theory of sampling describes and evaluates all errors involved in sampling of materials within a decision unit as well as methods for minimization of error to an acceptable level to meet SQC requirements.

Total sampling error (TSE): Error during any mass reduction stage that causes the measured concentration or characteristic of interest of the test portion to deviate from the true concentration or characteristic of interest of the decision unit. The major components of total sampling error are FSE, GSE, IDE, IEE, and IWE:

Fundamental sampling error (FSE): Imprecision error due to compositional heterogeneity.

Grouping and segregation error (GSE): Imprecision error due to distributional heterogeneity.

Increment delimitation error (IDE): Bias error due to incorrect shape of increments.

Total sampling error *(continued)*:

Increment extraction error (IEE): Bias error due to incorrect increment extraction.

Increment weighting error (IWE): Bias error due to inconsistent increment masses.

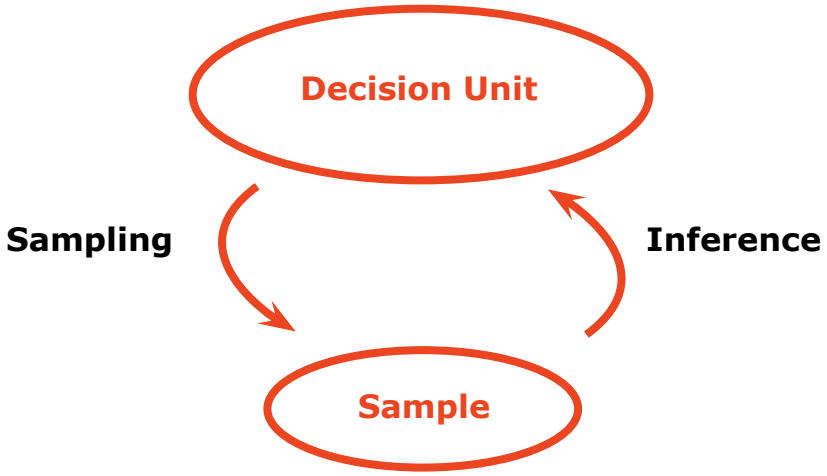
Acronym/Abbreviation	Definition
CH	Compositional heterogeneity
DH	Distributional heterogeneity
FSE	Fundamental sampling error
GEE	Global estimation error
GSE	Grouping and segregation error
IDE	Increment delimitation error
IEE	Increment extraction error
IWE	Increment weighting error
SQC	Sampling quality criteria
TOS	Theory of sampling
TSE	Total sampling error

Figure 1a. Sampling terms illustrated for tomatoes.

Figure 1b. Sampling terms illustrated for grain.

Representative Sample

A representative sample is one that can be used to answer a question(s) about a decision unit with an acceptable level of confidence. This requires a complete understanding of the sample quality criteria (**SQC**), developing the appropriate sampling protocol and including proper quality control to assess error. For example, if a mean concentration is desired for a decision unit, then the sample(s) collected should contain the same concentration of analyte as exists within the decision unit. A representative sample is one where both imprecision and bias are controlled to an acceptable level.



Imprecision error is controlled by collecting an appropriate mass and number of increments to address the compositional and distributional heterogeneity.

Bias error is controlled during sampling when every element in the decision unit has the same probability of being selected (equiprobable). Correctness is maintained when additional biases are not introduced during the sample preparation process. Because bias cannot be measured or completely eliminated, the term "sample correctness" is used when bias is controlled to a negligible level.

Therefore, a sample is said to be representative if the following conditions are met:

- It is correct (bias controlled to a negligible level), and
- It has a sufficiently small imprecision.

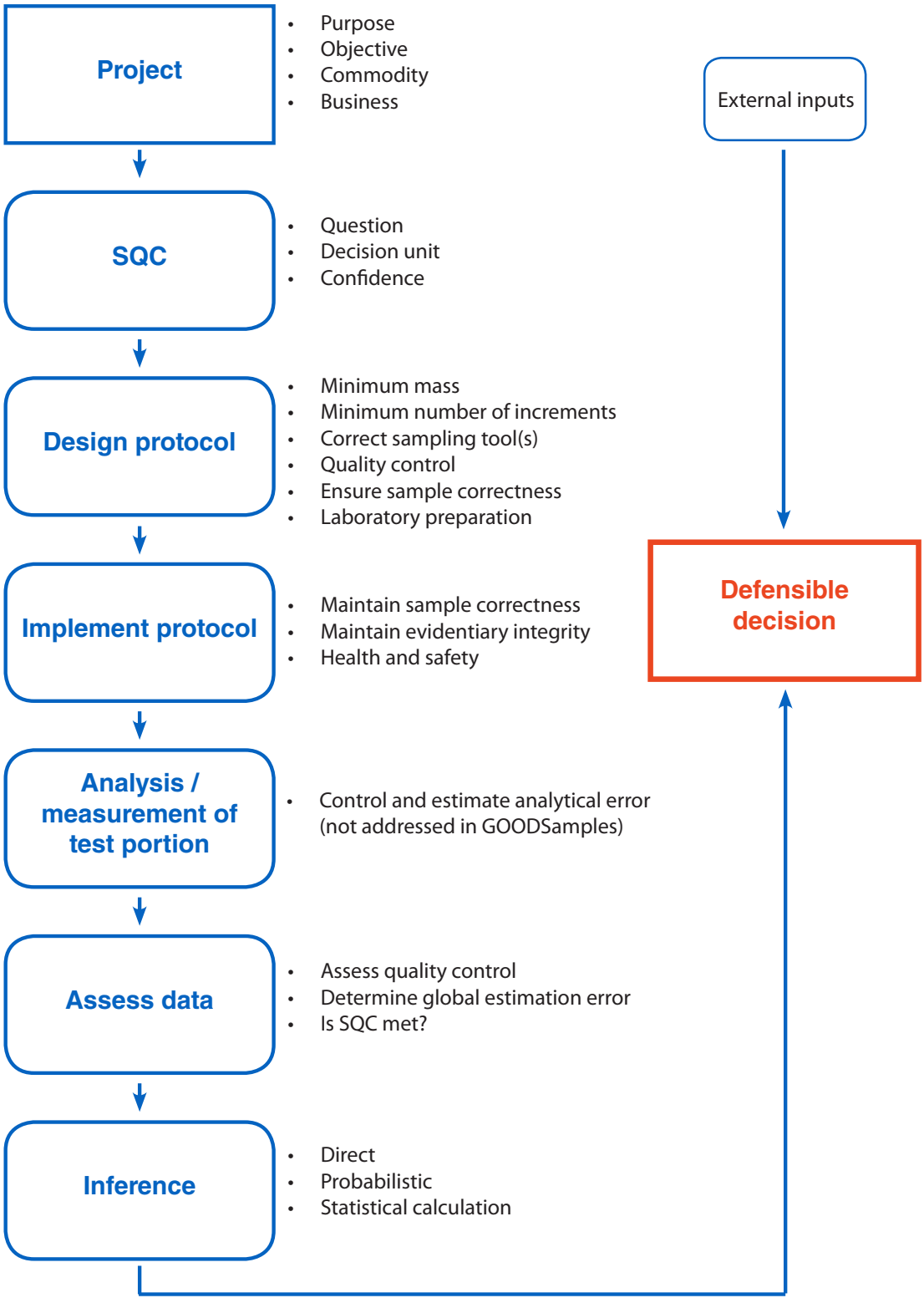


Figure 2. Flow chart for defensible decisions (overview of the *GOODSamples* approach).

MANAGEMENT CONSIDERATIONS

Management support is critical to implementing *GOODSamples*. This guidance describes a systematic process to obtain primary samples and test portions that provide a foundation for defensible decisions. Management must support good sampling practices as they ensure food or feed safety. Management should establish priorities related to sampling efforts, recognizing the direct relationship of sample collection and sample preparation to defensible decisions. Management support includes allocation of sufficient resources, including funding, staff, equipment, and training. The organization's administration and laboratory management should be fully aware of the theory of sampling principles when making decisions regarding allocation of budget and personnel. They should strive for high confidence in sampling and allocate the necessary resources.

Management should first determine objectives for sampling that dictate the type and quality of samples required to achieve the stated objectives with the most efficient use of resources. The management sampling plan defines the purpose and frequency for sampling, the types of food/feed commodities, and the firms/locations that may be sampled. Sampling protocols define the specific procedures for obtaining samples to meet the stated objectives in the management sampling plan.

Management must understand the types and sources of sampling errors. They must also understand techniques to mitigate sampling error in order that they provide the training and resources necessary. Management must recognize the importance of the proper tools and equipment. The availability of appropriate, well-maintained sampling tools must be a management priority, so that the sampling process is correct and sampling biases are negligible.

Management at all levels must have a solid understanding of the importance of sample preparation and handling upon the quality and defensibility of analytical data and impact on resulting decisions. Laboratory management support of sampling will be evidenced by adequate equipment and qualified human resources for laboratory sample preparation activities. Sample preparation facilities and activities are often the most

neglected area in laboratories, and are often not recognized as having a critical impact on data quality.

Management should provide adequate resources and support to assure scientifically valid sampling and sample preparation from point of collection through data analysis to

- Define the sampling purpose and objectives;
- Hire sufficient qualified sampling staff and laboratory staff;
- Delegate technical responsibilities and decisions;
- Ensure safety and health of employees;
- Determine the consequences of a wrong decision;
- Provide necessary tools, equipment, and supplies for sample collection;
- Conduct ongoing periodic training;
- Provide adequate facilities for handling of incoming laboratory samples to prevent cross contamination and to ensure analyte and evidentiary integrity;
- Provide and maintain necessary equipment for mass reduction, particle size reduction (comminution), sample mixing, and sample storage;
- Assure routine communication between sampling staff, laboratory staff, and management;
- Provide oversight, supervision, and periodic on-site observation;
- Implement a system for maintaining evidentiary integrity from collection of primary samples through data analysis and interpretation; and
- Implement a plan for project assessment, feedback and continuous improvement.

Coordination and Collaboration

Intra-organizational collaboration is essential. Planning may be for routine surveillance of products for specific adulterants and/or contaminants over a long range of time or planning could involve a time-sensitive matter that requires immediate coordination for a specific event. Coordination among all entities (e.g., management, sample collector,

program, laboratory) is necessary before undertaking sampling activities.

Establishing communication with all parties involved is critical. It is an advantage to agencies and industry to facilitate collection of representative samples because unrepresentative samples may result in the wrong decision (i.e., recall of a good lot). Collaboration among all stakeholders is the best way forward to accomplish collection of defensible samples.

Long-range planning may take the form of periodic meetings between compliance/inspection and laboratory managers to agree upon numbers of samples, identification of desired analytes, and logistics of submitting samples.

Short-term planning for a specific adverse event (e.g., violation investigation, special events, or foodborne outbreak) should also occur to ensure efficient operations. This planning could involve discussion of sufficient personnel, surge capacity, and expectations for turnaround time.

Coordination and collaboration may include

- Developing and communicating the sample quality criteria;
- Establishing how imprecision error will be calculated and reported;
- Establishing how sample correctness will be maintained;
- Establish how inference and resulting decisions will be made;
- Reserving samples for possible violation confirmation;
- Prioritizing and estimating the number of primary and laboratory samples that can be handled within an appropriate holding time;
- Determining what demographic information is to be collected;
- Implementing chain of custody;
- Organizing logistics for transporting laboratory samples, whether by commercial carrier, field staff delivery or other options;
- Identifying special handling needs upon laboratory sample receipt; and
- Planning for retention of a portion of the laboratory sample.

Note: Coordination with the servicing laboratory about reserved samples is critical. Issues to consider include at what step in the process should the reserved sample be generated and how it should be generated (e.g., prepared or unprepared, split or replicate).

Training

Management should identify personnel responsible for conducting and receiving training, and develop a written training plan that includes the reasoning behind the sampling plans and sampling protocols. Management should be trained on theory of sampling (**TOS**) prior to or concurrently with training of staff. Training is critical so that staff at all levels can make real-time decisions consistent with projects objectives when unanticipated conditions and situations are encountered. Training oversight should include a test of training effectiveness with both written testing and observation; documentation of initial and periodic training; periodic audits of quality assurance/quality control and feedback to all staff. The laboratory staff may need additional training including sample receipt, data review, compliance determinations, sample disposal, handling and storage of analytical samples, and timely reporting of results. Managers may choose to take advantage of external training opportunities. However, internal training must also occur to ensure coordination between management, inspection, sampling, and laboratory staff.

Training may include

- Scientific basis for testing (physical, chemical, and microbiological);
- Theory of sampling;
- Sample quality criteria;
- Maintaining sample correctness/analyte integrity;
- Quality control implementation and assessment;

- Maintaining evidentiary integrity;
- Documentation policies and practices;
- Overview of relevant regulations;
- Proper selection, use and maintenance of tools and equipment;
- Practical, hands-on application with representative products and sampling tools;
- Safety practices, including use of personal safety clothing and equipment (PPE);
- Collection of owner demographic and product trace-back data;
- Use of computers and data submission applications; and
- Procedures for conducting violations/compliance investigations.

Targeting

GOODSamples does not address the management prioritization or "targeting" of specific products or facilities for inspection but rather describes how the primary samples are taken once targets have been identified. Each organization targets its scope of work to meet the objectives in their areas of responsibility with the limited resources available. This is a local choice that affects allocation of staff and budget but is outside the scope of this guidance. Targeting, however, may affect the decision unit and, for this reason, management should be aware of sample quality criteria when targeting specific products or facilities.

Decision Making

In order to make defensible decisions, inferences must be based on a clear understanding of the errors inherent in the measurements, including sampling. It is the responsibility of those generating data to clearly identify the error associated with data. Protocols must be established that define actions (e.g., warnings or recalls) to be taken if analytical data indicates a food/feed hazard or non-conformance with established regulations.

These protocols may identify levels of actions relevant to the seriousness of the hazard and the level of confidence. An ongoing discussion about the quality of data can lead to improved procedures and greater confidence for decision makers. This document does not incorporate information beyond inference (e.g., economics or politics) in the decision-making process. Key steps in the decision-making process include the following:

1. Identifying the position/individual responsible for making decisions;
2. Designating a chain of command and alternate decision makers;
3. Establishing policy and protocols for making decisions; and
4. Incorporating the effect of global estimation error on inferences.

HEALTH AND SAFETY

It is the responsibility of the sampling staff to clearly identify all potential hazards before entering a sampling location. Hazards can include chemical, biological, radiological, electrical, and physical hazards, among others. Once all potential hazards have been identified, the appropriate personal protective equipment (**PPE**) must be selected and utilized at all times while on site. Personnel collecting samples should always be alert for possible dangerous conditions (e.g., poisonous materials or fumes, flammable or caustic chemicals, moving equipment, confined spaces). Additional health and safety considerations for the appropriate handling of the collected samples must be considered based on the nature of the sample. Appropriate warnings and precautions at the sampling site about the nature of any hazard(s) must be adhered to. A suitable container must be used to collect, transport, and store the samples to ensure any contamination cannot be unintentionally spread. In addition, assure tools are in good working order. Tools that jam or bind can be hazardous.

At a minimum, the sampling staff should observe the health and safety procedures and practices of the facility where sampling is being conducted. They should also consider the dual nature of PPE; it not only protects them from on-site hazards, it also protects the site and samples from possible contaminants introduced to the site by the sampling personnel or the sampling operation (e.g., residue on footwear, clothing, or hands from the previous site or sample). Cross contamination of one sample to the next, or from one site to the next, must be prevented. For example, diseases can be transferred from one swine breeding facility to another, so it may be necessary for the staff to shower and change clothes before entering the facility. Vehicles must be sufficiently clean to avoid transport of contaminants and disease among facilities. Also, changing gloves and washing hands should be considered before taking multiple samples from the same site. Keeping the outside of the sample container clean and uncontaminated is of utmost importance for downstream handling of the container. Special care is needed when handling samples to be tested for microbiological contamination (aseptic technique) and for drugs that may cause allergic

reactions in certain individuals (e.g., penicillin).

Sampling staff should be trained on health and safety before entering a site. It is management's responsibility to ensure that appropriate health and safety training is provided before any staff are deployed to the field.

SAFETY FIRST!

No sample is more important than personal safety

SAMPLING QUALITY CRITERIA

Sample quality criteria (**SQC**): A series of statements that clarify program technical and quality needs to support defensible decisions, including statement of the question to be answered, definition of the decision unit, and the desired confidence in the inference.

SQC applies from primary sample collection through laboratory analysis, including preparation of an analytical sample and selection of a test portion.

Simply stated, SQC is planning what you are going to do before you do it to ensure

- that stakeholders (e.g., management, sampling staff, laboratory staff, quality assurance staff, industry, consumers) agree on what needs to be done;
- that SQC is communicated to all relevant entities;
- the representativeness and applicability of analytical results;
- the optimal allocation of resources;
- that final data will be of sufficient quality to make a defensible decision;
- the equivalency of laboratory data; and,
- ultimately, the harmonization of decision making.

Sample quality criteria are determined based on responses to the following three questions:

(1) What is the question to be answered?

(2) What is the decision unit?

(3) What is the desired confidence in the inference?

The entire SQC must be developed before the sampling protocol (minimum mass, minimum number of increments, correct tools, maintaining analyte and evidentiary integrity and quality control) can be designed. Each element of the SQC is discussed in greater detail below. The "decision unit" is the material from which a sample is collected and to which an inference is made. The term is discussed in detail later in this section.

What is the Question to be Answered?

Question 1 can be broken further into separate questions:

a. What information is required?

- What is the analyte?
- What is the level of concern?

b. What type of data will be collected?

- Characteristic of the decision unit?
- Concentration of analyte(s) in the decision unit?

c. How is the inference from the sample to the decision unit going to be made?

- Direct inference (single result)?
- Probabilistic inference (single result)?
- Statistical inference (e.g., average of multiple results, confidence interval)?

Once the analyte and concentration of interest (or the characteristic) have been determined, the selection of tools and containers, sampling techniques, package types, preservation techniques, holding times, laboratory preparation equipment needs, detectability, and PPE can be identified. The range of possibilities is almost limitless when all the combinations are considered, especially for multiple analytes from the same primary sample. Sorting through all the possible combinations requires communication with laboratory staff, sampling staff, management, and other professionals (e.g., collection site personnel, shippers, receivers). The goal is to assure that analyte integrity is maintained from primary sample collection through analysis. If the potential exists that analyte integrity may be compromised, changes must be made (different tools, shorter holding times, preservation) or quality control events (e.g., blanks) added to monitor analyte integrity.

Inference to a single decision unit can be made in a variety of ways. Where a single analytical result is used to estimate the concentration of analyte of interest in the decision unit, there can be a either a direct or a probabilistic inference. Alternatively, some type of statistical calculation can be used (e.g., average, 95% upper confidence limit) to make an inference to the decision unit. The latter type of inference will require multiple analyses.

When the number of decision units is large, it may not be possible to sample and make inference to each of them individually. In these situations, a percentage of decisions units may be sampled and inference made to the unsampled decision units. The process of making inference from sampled decision units to all decision units is a different approach from the process of making inference to a single decision unit.

There are three common ways to make inference to a decision unit: direct, probabilistic, and based on statistical calculations.

- *Direct inference:* When the entire decision unit is selected and analyzed in its entirety, inference is made directly from the test result to the decision unit. In this case, the only source of error would be analytical error.

- *Probabilistic inference:* When the entire decision unit cannot be selected and analyzed in its entirety, inference is made directly from the test result to the decision unit without performing any statistical calculations. With probabilistic inference, confidence is evaluated based on the magnitude of the global estimation error (**GEE**). Global estimation error is estimated from process-wide quality control samples.

- *Inference based on statistical calculations:* If multiple results are available, inference can be based on statistical calculations. In cases where the confidence is not included in the calculation, the GEE must be determined.

What is the Decision Unit?

The decision unit is the material from which the primary sample(s) is collected and to which the inference(s) is made. There may be one decision unit or multiple decision units. The term "decision unit" is used as opposed to "population" or "lot" to emphasize that this is the material from which the sample is collected, to which the analytical results apply, and to which inference is made. For instance, a truckload of cantaloupes may be the population of interest. However, if the objective is that no cantaloupe can be contaminated with *Listeria monocytogenes*, the decision unit must be the individu-

al cantaloupe and there are many of them. Definition of the decision unit is the most critical, yet most often overlooked, aspect of a sampling design. Regulations may dictate the choice of decision units. Specific targeting of certain commodities or businesses may determine the choice of the decision unit. The decision unit must be based on regulatory or scientific criteria. It cannot be selected based on sampling convenience. Decision units need to be unambiguous and clearly understood by all parties. There can be no doubt when a material is part of a decision unit and when it is not.

The decision unit in the laboratory is still the original decision unit that the primary sample is intended to represent; however, the laboratory is restricted to the laboratory sample as they receive it. Their responsibility is to ensure that all test portions represent the laboratory sample and thus, the original decision unit.

A requirement for an unbiased inference is that the decision unit must be accessible. If it is not, then effort must be made to make it accessible for the purpose of sampling. It is incorrect to make the accessible portion of the decision unit the "new" decision unit because the goals of the project will be compromised, leading to incorrect decisions and inequivalent data. The decision unit must be communicated and respected throughout the entire measurement process. It cannot be changed for convenience. For example, if the objective is to know the average concentration of all the drums of honey in a warehouse, then all of the drums need to be accessible and randomly sampled.

***Example*:** Fifteen pallets of 100 containers each are present. The scale of information desired drives the decision unit. In the case of food/feed safety, is the public protected if

- The average value of all 15 pallets is found in the safe range? If so, then all 15 pallets comprise the decision unit (1 decision unit).
- Or must each pallet individually be safe? Then, each of the 15 pallets is the decision unit (15 total decision units).
- Or must each container on every pallet be safe? Each container in all 15 pallets is the decision unit (1,500 decision units).

What is the Required Confidence in the Inference?

If the legal, health, or economic risk related to an incorrect decision is high, more confidence in the analytical result is required. To achieve more confidence in the inference, error must be controlled to a greater extent. Keep in mind that better sampling does not necessarily mean more primary samples. During the development of the SQC, the minimum confidence requirement for any type of inference must be established.

- *Direct inference:* Because the entire decision unit is selected and analyzed in its entirety, there is no sampling or sample processing error and confidence is a function of analytical or measurement error.

- *Probabilistic inference:* With probabilistic inference, because only a portion of the decision unit is selected or analyzed, errors related to the mass reduction stage(s) exist in addition to the analytical or measurement error. Probabilistic inference requires sample correctness and an estimate of the GEE to determine confidence. In no case should the tolerable GEE be specified at greater than 35% relative standard deviation.

- *Inference based on statistical calculations:* Statistical calculations require correct (unbiased) replicate samples, each meeting the requirements of probabilistic inference. The statistical calculation chosen will dictate the number of replicate samples. With inference based on statistical calculations, the specific calculation and confidence level must be defined. Examples of a statistical calculation are an average of several measurements or a confidence interval of the mean. For a confidence interval of the mean, the GEE is incorporated into the calculation and must be based on implementation of *GOODSamples*.

Summary of SQC

The SQC is a living process that allows feedback from all levels when circumstances or original assumptions change. The development of SQC involves input from all areas of the organization. Targeting may be part of the management input into SQC.

SQC can occur at different levels:

- **Program-wide SQC:** When SQC is a programmatic process, the SQC can be used to develop standard operating procedures (**SOPs**) for specific recurring sampling events. However, even with a well-written SOP, there may be conditions at the sampling site that cannot be anticipated and sampling staff will need to use their judgment and experience. When these situations arise, if the actions taken

are through the context of SQC, equivalent and defensible outcomes are likely. Without the context of the SQC, equivalency and defensibility are at the whim of judgment or intuition, not of a plan with known confidence. It is essential that the sampling staff understand the SQC inputs, so they can make proper on-site level decisions that may result in deviations from the SOP. Proper documentation provides a record of onsite observations, activities, and decisions so that they can be taken into account during data review and assessment. The laboratory needs to know the level of programmatic concern for each analyte. For example, is testing requested to verify a label guarantee or for residue contaminants, and at what concentration is programmatic action initiated or triggered? This will determine the precautions the laboratory needs to employ while handling the laboratory sample. When performing residue or contamination level analysis, for example, all preparation equipment must be free of any traces of that analyte, especially if the same equipment is used to prepare samples that contain the analyte. The equipment must also be able to be readily cleaned so that subsequent materials are not cross-contaminated. The laboratory must also have a good understanding of the total composition of the laboratory sample submitted due to potential interactions or analytical interferences that the other ingredients may have on the analyte of concern, and to assess sample processing needs.

- **Sampling-level SQC:** There are situations that cannot be anticipated by routine SOPs or even the best planning. It is especially critical in these situations that the SQC process be developed and applied at the sampling location to ensure defensibility of the data. For example, during an investigation, a suspicious activity or product is observed and a sample is warranted to investigate the observation. Other examples may be an outbreak, specific complaints, or spills. Coordination with the laboratory and management is especially critical in these situations.

- **Laboratory-level SQC:** For routine laboratory samples, the laboratory should have participated in the programmatic SQC process and laboratory concerns would have been dealt with at that time. Laboratory-specific SQC is implemented when the laboratory receives a sample that has no program-wide or sampling-level SQC. In these cases, the laboratory needs to develop SQC before sample preparation and analysis that address the same concerns as expressed above in the programmatic SQC. Laboratory-level SQC is also implemented when unanticipated circumstances occur or observations are made that were not included in programmatic SQC.

For developing a sampling protocol, the identification of the decision unit should be

completed in the SQC section and then what remains are issues related to how to randomly access the decision unit and retrieve the correct number, mass, and shape of increments. Random access may require the physical movement of the decision unit. Tools that provide access to all portions of the decision unit may address accessibility issues. The easiest and best method to collect a sample is if sampling can be timed to occur when the material is in motion; therefore, whenever possible, sampling should be coordinated with the movement of the material to be sampled.

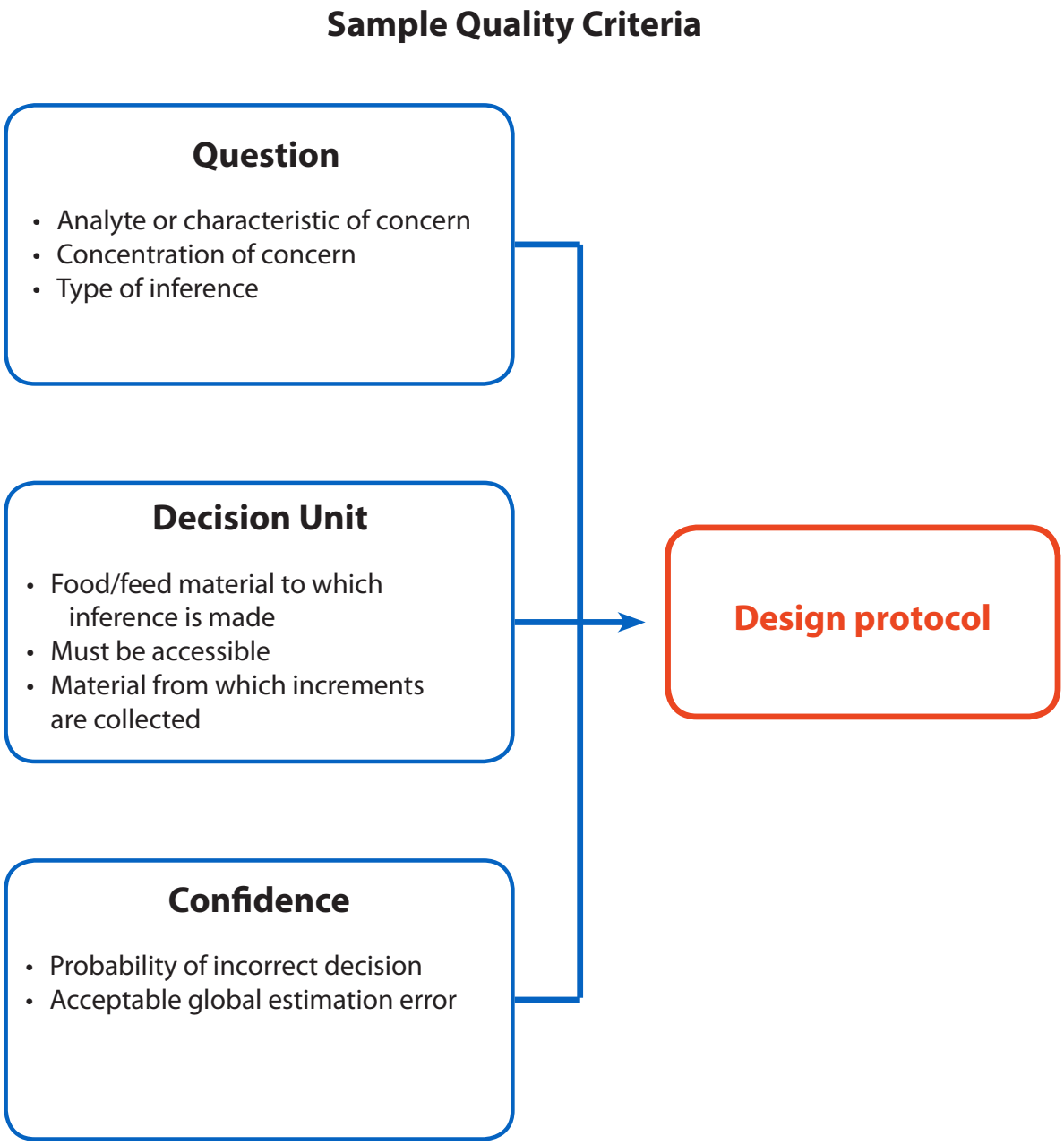


Figure 3. Sample quality criteria.

MATERIAL PROPERTIES

Material properties include those properties of the material that comprises the decision unit that must be considered when developing sampling protocols. Material properties include material elements and heterogeneity. Each must be carefully considered in the design of the sampling protocol.

Material Elements

Finite Element Materials. "Finite element material" is a term utilized in this document to describe a decision unit that consists of a finite number of discernible elements. These elements can be individually identified and selected at random from the decision unit. The important concept for sampling of finite element materials is the ability to identify an individual element at random and then to select that specific element from the decision unit. These finite elements are combined to form a primary sample. Finite elements can be naturally occurring or the result of packaging (e.g., bags of spinach or cans of tuna fish).

Finite element material sampling strategies can be devised for any confidence level. If more elements are collected, there will be greater confidence in the inference regarding the decision unit. Finite element material sampling strategies are the easiest to develop and implement, as the only variable is the number of elements selected.

Infinite Element Materials. "Infinite element material" is a term utilized in this document to describe a decision unit that consists of a practically infinite number of indiscernible elements. The individual elements cannot be individually identified prior to sampling nor can they be collected individually. Some examples are flour, soil, grain, leaves, water, and salsa. The important concept for sampling an infinite element material is the notion that elements can neither be individually identified nor individually selected at random. Sampling of decision units composed of infinite element materials is more complex than sampling those composed of finite element materials.

Because infinite element materials do not consist of individually identifiable elements, groups of elements must be collected from the decision unit. These groups of elements that are removed by the single operation of a sampling tool are called "increments." An example may be a scoop of dirt, a probe of grain, or a spatula of ground feed. The selection, removal, and recombination of increments form the basis of any mass reduction throughout the process (e.g., primary sample, analytical sample, test portion). Key considerations in the sampling of infinite element materials are the final mass, the number of increments, and the shape of the increments, which are not a consideration for finite element strategies.

Finite Versus Infinite Element Materials. When sampling finite element materials, the only consideration is the number of random elements (i.e., increments) collected. The physical properties of the finite element material are fixed and therefore, there is no ability to control the mass and shape of the increment. However, with infinite element materials, the mass, shape, and number of the increments must be controlled. If the mass and shape of the increments are not properly controlled, errors will be introduced in the sample collection process that will make the sample unrepresentative. In many cases, sampling protocols for infinite element materials are inappropriately based on finite element sampling protocols.

Example of finite element material: A decision unit comprises a pallet of cans of tuna fish. In this case, the cans are finite elements. Individual cans of tuna fish can be identified and collected at random throughout the pallet.

Example of infinite element material: A decision unit is a bulk bag of flour. In this case, it is impossible to locate and collect individual particles (elements) of flour. Therefore, groups of particles, or increments, must be collected at random with a sampling tool from throughout the bulk bag.

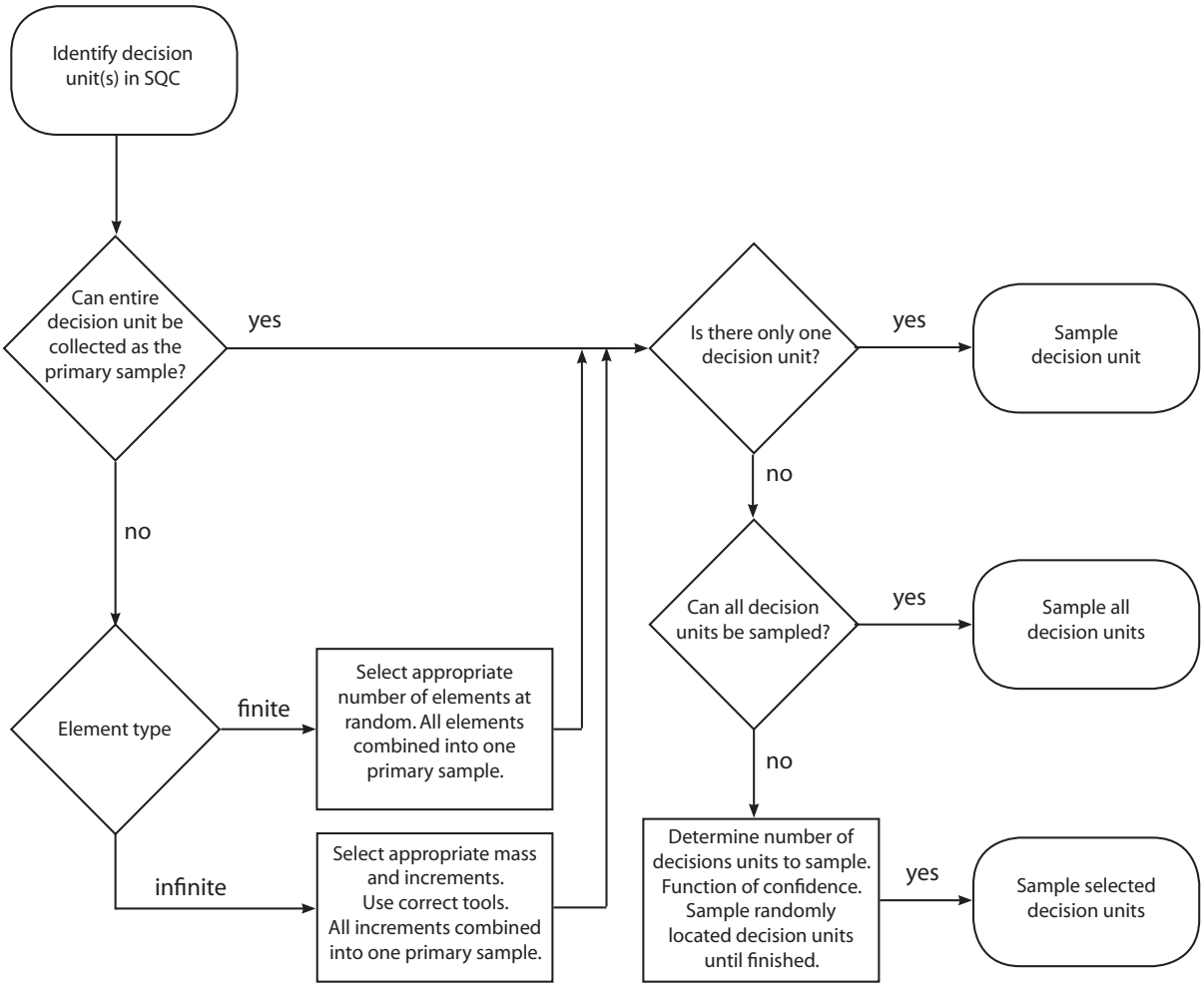


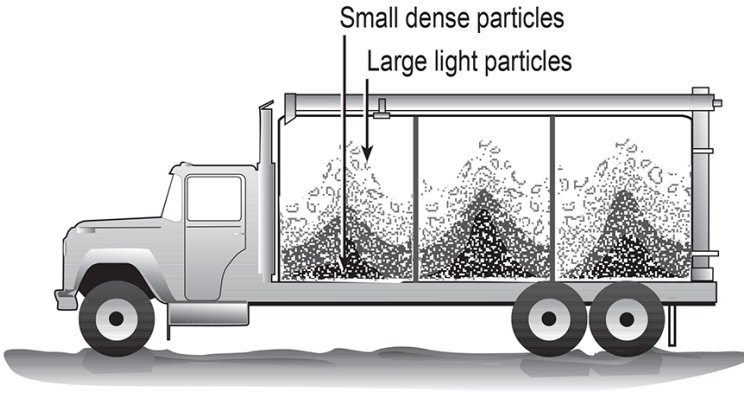
Figure 4. Flow chart for sampling of finite and infinite materials.

Heterogeneity

Design of sampling protocols must incorporate the concept of heterogeneity. There are two types of heterogeneity: compositional (referred to in some literature as constitutional) and distributional (referred to in some literature as spatial or temporal). The magnitude and nature of compositional and distributional heterogeneity are unique to every material. The magnitude of the compositional and distributional heterogeneity dictates the sampling effort.

Compositional Heterogeneity. Compositional heterogeneity exists when the individual elements (particles for solids; molecules for liquids and gases) that make up the decision unit exhibit differing concentrations of the analyte of interest. Compositional heterogeneity always exists to some degree and cannot be altered without comminution. Mixing has no effect on compositional heterogeneity.

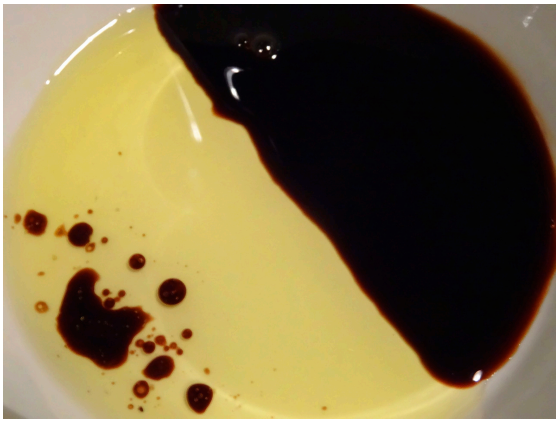
Distributional Heterogeneity. Distributional heterogeneity results from non-random distribution (spatial or temporal) of elements within the decision unit. A good example of distributional heterogeneity is the settling of small, dense fines to the bottom of a container. Distributional heterogeneity always exists to some degree, and it is altered with physical manipulation of the material (e.g., vibration causing segregation, mixing).



Reproduced with permission from Joseph P. Harner III, et al., *Avoiding Drug Carryover During Feed Processing and Delivery*, Kansas State University, March 1996.

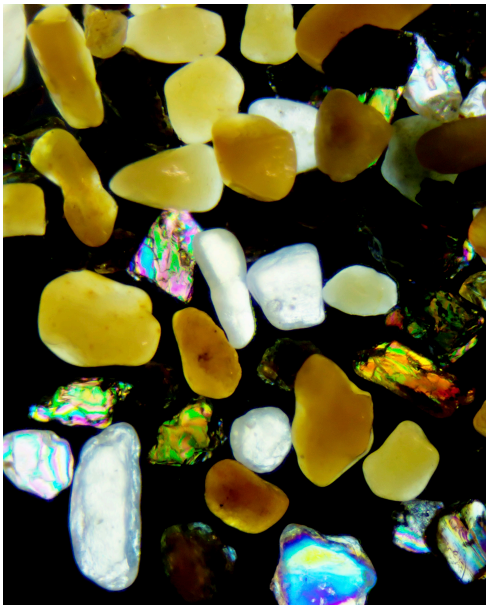
These pills represent a finite element material with compositional heterogeneity.

This granola illustrates a solid infinite element material with both compositional and distributional heterogeneity.



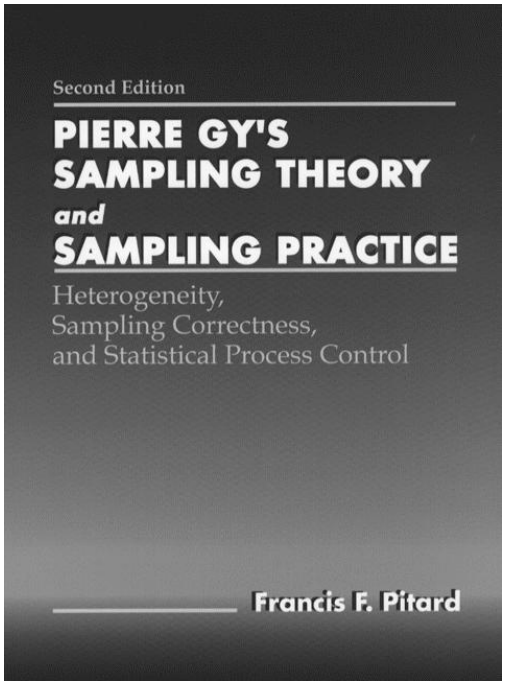
The oil and vinegar illustrates a liquid infinite element material with both compositional and distributional heterogeneity.

Sand may appear uniform from a distance, but when examined with a microscope, it possesses large compositional heterogeneity.



THEORY OF SAMPLING

A critical aspect of sample collection is an understanding of the Theory of Sampling (**TOS**) and what makes a representative sample. Understanding the TOS is important for management, quality assurance staff, sampling staff, laboratory analysts, and data users. Without this knowledge, it is impossible to develop effective sampling protocols, manage resources, or make defensible decisions.



The TOS covers both finite and infinite element materials; however, the primary focus is infinite element materials. The sampling of infinite element materials is often not properly addressed in food and feed sampling guidance, which has typically been based on finite element sampling.

Introduction to Sampling Errors

Fundamental Sampling Error. Fundamental sampling error (**FSE**) results from compositional heterogeneity, which is controlled through the collection of sufficient sample mass (mass is indirectly proportional to error). The larger the compositional heterogeneity, the greater the sample mass that must be collected. It is called FSE because it is the error that exists even if all other sampling errors are perfectly controlled. There are various formulas to estimate the FSE. A simplified equation that relates the FSE, particle size, and sample mass follows:

$$FSE^2 = \frac{Cd^3}{m_s}$$

where FSE = fundamental sampling error; C = sampling constant, d = diameter of the largest particles (cm), and m_s = mass of the sample (s) collected (g). The sampling constant (C) is unique for each type of material.

The mass required to control the FSE to a specific level for any material can be determined by using the fundamental sampling error equation. The FSE equation requires substantial knowledge about the material being sampled.

Grouping and Segregation Error. Grouping and segregation error (**GSE**) results from distributional heterogeneity, which is controlled through the collection of a "sufficient number" of random increments. The larger the distributional heterogeneity, the greater the number of increments that must be selected and combined for the sample. The root cause of GSE is the inability to collect the individual elements one at a time at random. For infinite element materials, groups of elements (increments) must be selected. An increment may contain thousands of elements but it is still located at one specific random location. As more and more random increments are included in the sample, the GSE decreases. If the mass required to control the FSE could be collected one element at a time at random, GSE would not exist.

There is no simple calculation to determine the number of increments to collect, but three basic approaches can be used to reduce the GSE—reduce the FSE, increase the number of increments that make up the sample, and reduce the distributional heterogeneity of the material.

(1) ***Reduce the Fundamental Sampling Error:*** The GSE is linked to the FSE; thus, reducing FSE will reduce GSE. It is intuitive that as more and more mass is collected, sampling errors will be reduced to the point where if all the mass from the decision unit were collected as the sample, there would be zero sampling error.

(2) ***Increase the Number of Increments:*** Increasing the number of properly collected increments is the easiest method to reduce the GSE. There is no magic number of increments. The number of increments is completely dependent on the degree of distributional heterogeneity of the material within the decision unit. Thus, there exist materials and SQC where 10 increments are sufficient, and there exist materials and SQC where 100 or more increments are required. The number of increments is never based on what is easy to collect but rather on the number required to minimize the GSE to the level to achieve the SQC. It would be inappropriate to always collect an identical number of increments (e.g., 10, 30, or 100) as standard practice for all sampling situations.

(3) ***Reduce the Distributional Heterogeneity of the Material:*** Reduction of distributional heterogeneity is an effective method to reduce GSE. A common technique to achieve this is mixing; however, mixing is only effective when the material is composed of elements that have a relatively uniform shape, size, and density. If elements are not uniform, mixing and stirring of the material does not significantly decrease distributional heterogeneity and may, in fact, increase it. It is

typically impossible to mix the entire decision unit at the primary sampling site unless the decision unit is relatively small. In the laboratory, it may be desirable to mix the analytical sample before removing the test portion. This is usually only beneficial if (1) some type of sample preparation (e.g., comminution) has previously occurred to reduce particles to a uniform size and shape, and (2) an appropriate mixing technique will be employed.

It is critical that increments be collected at random, using the correct tools, from throughout the entire decision unit. A variety of forms of random increment selections exist, each with pros and cons. However, as long as a sufficient number of increments are selected at random, the sample will be representative and inference can be made.

Illustration of both compositional and distributional heterogeneity.

Sample Correctness. Sample correctness is the condition achieved when bias errors are controlled to a negligible level. Sample correctness is achieved when selection of elements at increment locations is equiprobable, and it is controlled by proper use of a correctly designed sampling tool. Once sample correctness is achieved with the primary sample, it must be maintained in subsequent stages all the way to generation of the test portion.

Sampling bias results from the following sources.

- Increment delimitation error (**IDE**): IDE results when the selection of elements at increment locations is not equiprobable. IDE is controlled through correct increment shape (see Table 1).

- Increment extraction error (**IEE**): IEE results when the increment is not collected and/or removed in its entirety. IEE is controlled through the proper use of the correctly designed sampling tool.

- Increment weighting error (**IWE**): IWE results when the increments are not the same size or proportionate. IWE is common with scoops when volume is not controlled. IWE also occurs when the sampling tool does not fill consistently, as with some types of probes.

Sample correctness ensures there is no preferential selection or avoidance of elements based on, for example, size (e.g., too big for the sampling tool), location (e.g., hard-to-reach area), or chemical property (adheres to sampling tool).

Sample correctness is achieved when the bias error from these sources is controlled to a negligible level. Although analytical biases are routinely estimated for analytical measurements, sampling bias is very difficult, if not impossible, to estimate. Sampling bias is inconsistent; therefore, conventional bias correction techniques to estimate and correct bias cannot be used.

Total Sampling Error. There are no established rules for the magnitude of GEE that can be tolerated in a sampling protocol but it should not exceed 35%. The error threshold is determined by the error that can be tolerated in the final result and is related to the desired confidence that a correct inference was made (SQC). When designing sampling and analytical protocols, the GEE must be estimated so that confidence in a final result can be determined.

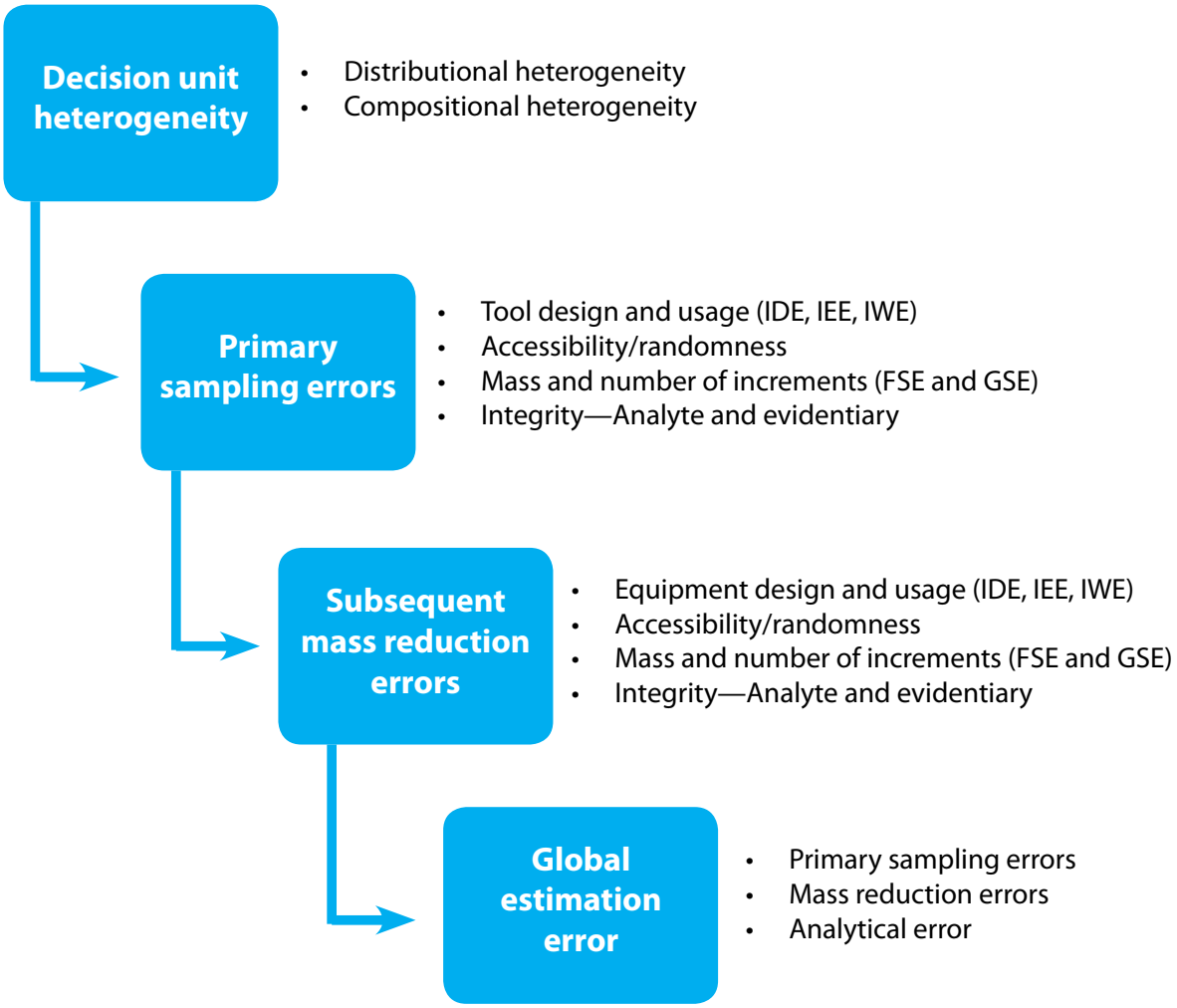


Figure 5. Contributors to global estimation error (GEE). The calculation for GEE is possible only when bias errors are negligible.

Summation of Errors. How GEE will be estimated and reported needs to be established. Global estimation error includes all imprecision and bias errors. Sampling bias is typically controlled to a negligible level (sample correctness) rather than measured and reported. Analytical bias is typically measured but not included in the analytical error. Imprecision error components of GEE are FSE, GSE, and analytical error. The precision error components of GEE include the precision error from each sampling (mass reduction) stage and analysis. There are typically many precision errors in the measurement process. Precision errors are independent and do not add up directly but propagate according to the following formula:

$$Global\ Estimation\ Error = \sqrt{a^2 + b^2 + c^2 + ... + n^2}$$

where a, b, c, …, n are individual imprecision errors from each sampling (mass reduction) stage and analysis. FSE and GSE are the major components of imprecision sampling error. Total imprecision sampling error, including sampling error in the laboratory, plus analytical imprecision error equals GEE.

Errors that are relatively small compared with other errors have little effect on the GEE. Mitigating the largest errors will have the most dramatic effect on GEE. For example, consider a case where the FSE is 35% and the GSE is 10%. The overall precision error is 36%:

$$Global\ \ Estimation\ \ Error = \sqrt{0.35^2 + 0.1^2} = 0.36$$

Reducing the GSE further would not reduce the overall precision error significantly. If the GSE is reduced to a point where it is small compared with the FSE error (<25 to 30% of the total error), further reduction in error can only be achieved by increasing mass.

In some literature, the word "uncertainty" is used in place of the word "error," especially with analytical error. In *GOODSamples*, the word "error" is used exclusively. A sampling protocol for defensible decisions is independent of this terminology.

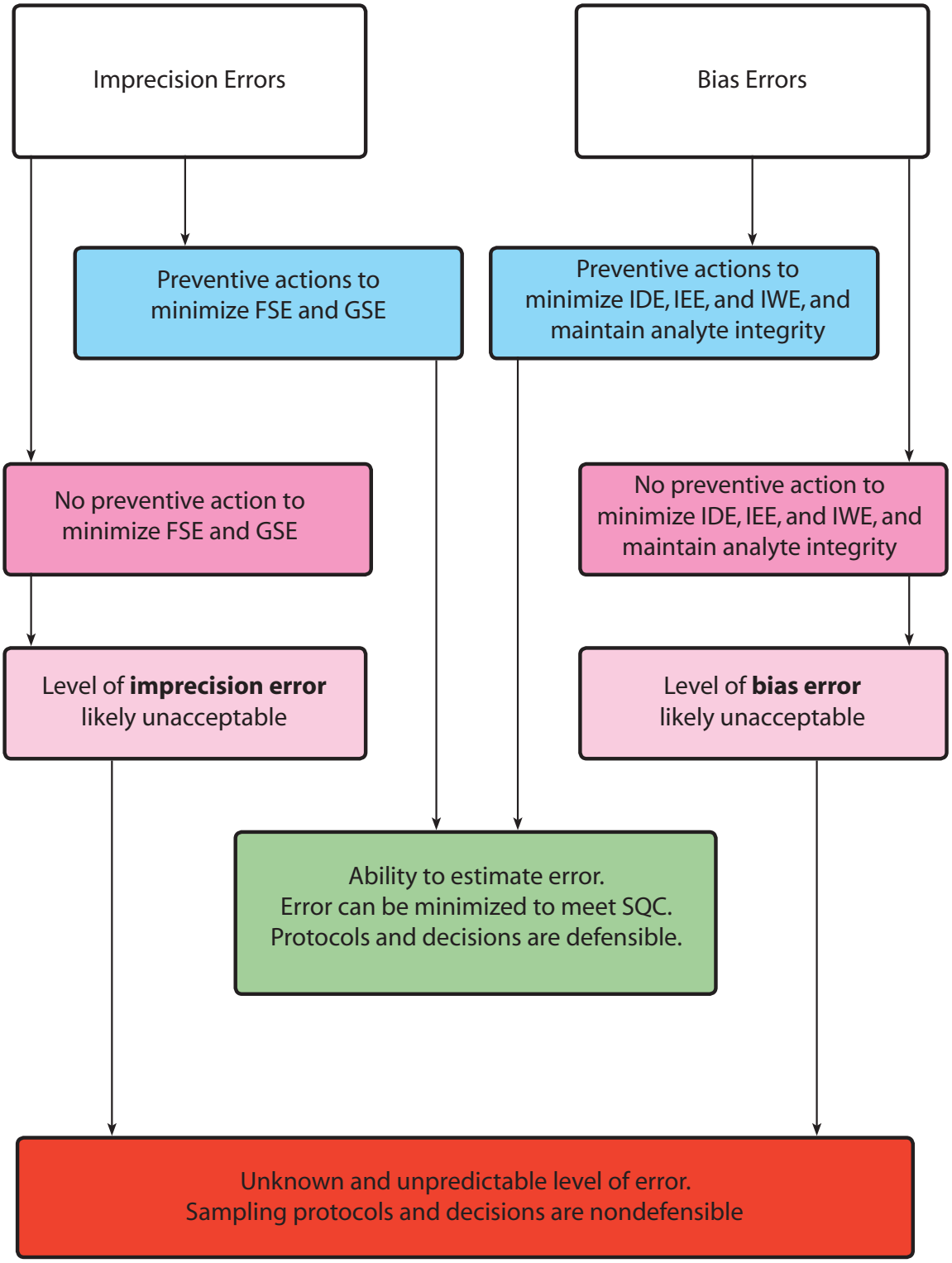


Figure 6. Sampling errors and their relationship to defensible decisions.

Primary Sampling Issues

Collection of sufficient mass to control FSE is generally not a challenge in primary sampling. The challenge in primary sampling is typically accessibility to the material in the decision unit and the collection of enough random increments from throughout the entire decision unit to control GSE. This is especially true for sampling a pile, truckload, or container presented as a decision unit when it is not possible to move the material to access all potential random locations. Therefore, the best time to sample a solid is when it is in motion (e.g., as the pile is created, the container is filled, or the material is moved from one location to another) because the moving process allows easy access to the entire decision unit and the GSE can be adequately controlled. It follows that, whenever possible, sampling should be scheduled when materials are being moved.

Laboratory Subsampling Issues

Collection of multiple increments to control GSE is generally not a challenge in laboratory subsampling. The challenge in laboratory subsampling is typically the collection of sufficient mass from the analytical sample to control the FSE. This is especially true when a small mass is desired as a test portion. Control of FSE is accomplished by comminuting the entire laboratory sample to a fine particle size or by keeping the test portion mass large. When comminuting to control the FSE, it is advisable to use a comminuting method that yields particles of a uniform shape and size. A uniform shape and size of particles will minimize the potential for additional distributional heterogeneity and therefore reduce the GSE. The comminuting process always produces an infinite element material, even if the original material was a finite element material.

Special Applications of the Theory of Sampling

Sampling of Surfaces. The materials sampled from surfaces typically have a small element size (e.g., microbes, dust, liquids), so FSE is small but GSE can be large due to distributional heterogeneity. A major source of confusion in surface sampling is identification of the decision unit (e.g., is the entire surface area in the factory a decision unit, or is each 100 cm^2 a decision unit, with millions of decision units in the factory?). If the decision unit is small enough that it can be collected in its entirety with a single wipe, then there is no sampling error, providing the swiping is effective. If the decision unit is so large that it cannot be collected in its entirety, then wipes from multiple areas (increments) may be combined to form the sample. This could be accomplished with one or more wipes.

Sampling of Containerized Liquids. Liquids present different sampling challenges than solids, with two main differences. One is the small size of the "elements" (essentially molecules and potentially small suspended particles) that make up the decision unit. For most liquids, the FSE is very small because the "particle size" is at the molecular level. However, FSE can still exist in liquids because of suspended particulate material. The other difference is the potential change of the decision unit over time. This temporal change is another aspect of distributional heterogeneity. Distributional heterogeneity may result from chemical properties, temperature graduations, viscosity differences, and a host of other chemical and physical phenomena.

For liquids, the mass or volume typically collected is enough to control FSE to a point where it is not significant. However, increments still need to be collected to control the GSE. Depending upon the SQC, the increments of a liquid may need to be collected over time as well as space. The following statements compare and contrast the differences with respect to temporal sampling of liquids:

(1) A representative sample of a static container of a liquid would need a sufficient number of increments collected at random throughout the entire container. A unique case of sampling a containerized liquid at a single point in time is when a single complete "core" can be taken through the entire depth of the liquid. In this case, the singe core can represent the entire container of liquid.

(2) A representative sample of a container of liquid that is part of a continuous process where the decision unit is determined to be the tank over a 24-hour production cycle would require a sufficient number of random increments throughout the tank and over the 24-hour period.

SAMPLE CORRECTNESS AND SAMPLING TOOLS

Tools used for collection of increments and equipment used for mass reduction must follow the requirements of sample correctness. Two types of tools will be discussed: sampling devices and splitting techniques. The purpose of a sampling tool is to provide access to and remove correct increments from a decision unit. The purpose of a splitting technique is to provide equal portions of the larger mass.

Sample Correctness

Sample correctness is a critical component of the collection of a representative sample and is necessary to achieve an unbiased test portion. Sample correctness requires (1) that all elements within a decision unit have an equiprobable chance of selection during the sampling process; and (2) that increments are proportionate. It is then critical that sample correctness is maintained during handling and preparation. In this chapter, these two requirements necessary for the collection of a correct sample are discussed. Maintaining correctness after primary sample collection is discussed in Evidentiary (and Analyte) Integrity and again in Laboratory Sampling and Preparation.

Equiprobable selection of elements has two components. The first is the ideal shape of the increment and the second is the complete extraction of the increment. Tools that do not obey the rules of equiprobable selection will introduce bias into the sampling operation. By its very nature, this bias is not consistent and cannot be measured. It is therefore essential that correct sampling be implemented to mitigate the impact of bias.

To ensure sample correctness, elements cannot be preferentially included or excluded based on size, location, or any other chemical or physical property. Some examples of bias include (1) over- or underrepresentation of elements based on size, chemical property, physical property, and so on (Figures 7 and 8), and (2) an analyte that adheres to the tool surfaces and, therefore, is not included in the primary sample, analytical sample, or test portion.

Errors that need attention when addressing correctness are increment delimitation error (**IDE**), increment extraction error (**IEE**), and increment weighting error (**IWE**). Increment delimitation is a concept that requires a proper pairing of the geometry of a tool with the dimension of the decision unit. In other words, the dimension (or shape) of the decision unit determines the ideal increment shape and dictates the design of the sampling tool. Increment extraction is the ability to remove the ideal increment shape in

its entirety. For the purposes of sampling, there are four dimensions of material: zero-, one-, two-, and three-dimensional (Table 1). Finally, it should be obvious that increments must be proportionate or IWE is introduced. An example of how IWE can be introduced is with the incomplete filling of an otherwise correct tool, making increments inconsistent in size and therefore, not proportionate.

Table 1.

Sampling dimensions	Correct increment shape	Tool example
Zero	Entire element	Hand
One	Cut or slice	Stream cutter, square spatula
Two	Cylinder	Coring device
Three	Sphere	Does not exist

It is important that all increments be the same size to avoid introducing another error, increment weighting error (IWE).

Sampling Tools

One criterion to achieve sample correctness is proper increment shape. The proper increment shape depends upon the dimension of the decision unit. A zero-dimensional decision unit is one where the material is finite and the time or spatial order is unimportant. Typically, elements are hand selected from a zero-dimensional decision unit and shape is not relevant because an increment consists of an individual element. Examples of a zero-dimensional decision unit might be cantaloupe or prepackaged goods such as frozen pizza.

A one-dimensional decision unit is one where one dimension is large compared with the other two. Examples are a material flowing along a conveyer belt or a falling stream. The correct increment shape is a slice across the entire width and depth (two dimensions) of the material. One-dimensional decision units are common in manufacturing environments.

A two-dimensional decision unit is one where two dimensions are large compared with the third. An example would be the top several centimeters of soil in an agricultural field or a vertical cylinder, such as a drum. A truck or railcar could be considered two-dimensional if cores can be taken through the *entire* depth. The correct increment shape for a two-dimensional decision unit is a cylinder core through the entire depth of the decision unit.

A three-dimensional decision unit is one in which all dimensions are large (no dimen-

sion can be sampled in its entire depth, width, and length with a sampling tool). Examples are a pile or silo. The correct increment shape for a three-dimensional decision unit is a sphere. Because no sampling device is available to collect spheres, errors are unavoidable.

The dimensions of the decision unit are directly related to the accessibility of the material in the decision unit. Inaccessible material cannot be sampled. Zero- and one-dimensional units are generally accessible and therefore the easiest dimension decision units to sample. Two-dimensional decision units can also be relatively easy to sample with appropriate tools. The sampling of three-dimensional decision units is very difficult and error prone, and should therefore be avoided wherever possible. The dimension of a decision unit can be changed by movement of the material. When a three-dimensional decision unit is encountered, it is advisable to flatten the material to convert it a two-dimensional decision unit or to move it to convert it to a one-dimensional decision unit. It is always advisable to coordinate sampling with the movement of material (for another

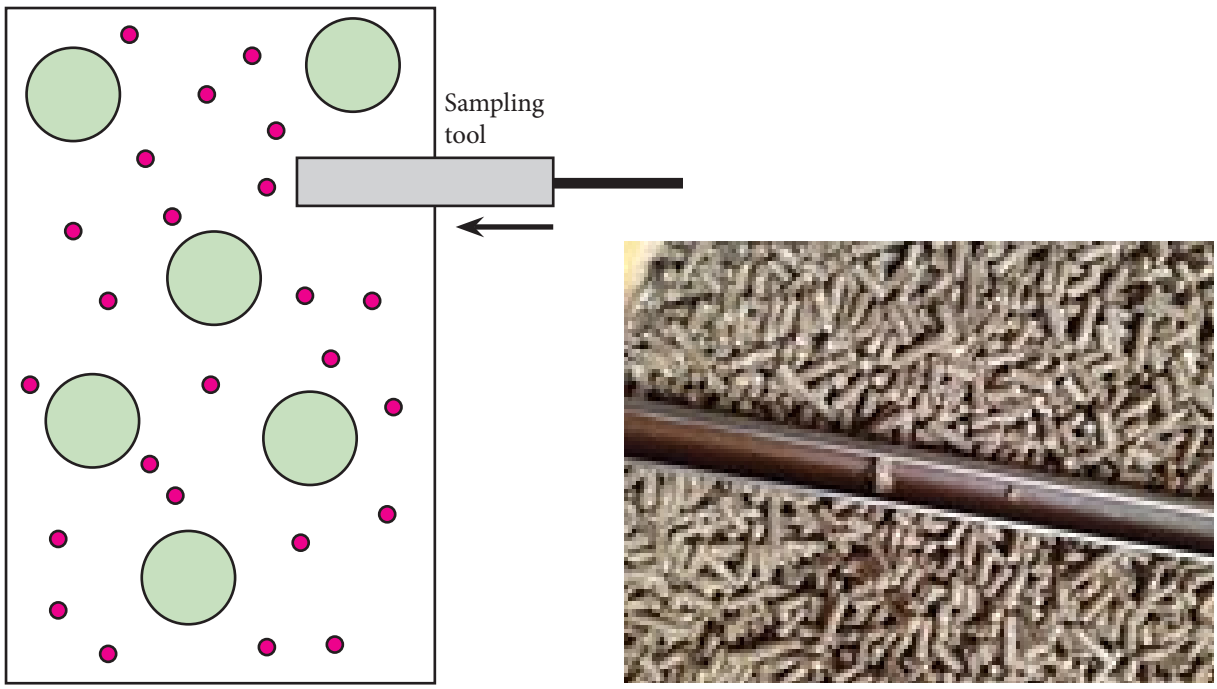


Figure 7. Example where the tool is too small to correctly capture the largest particles.

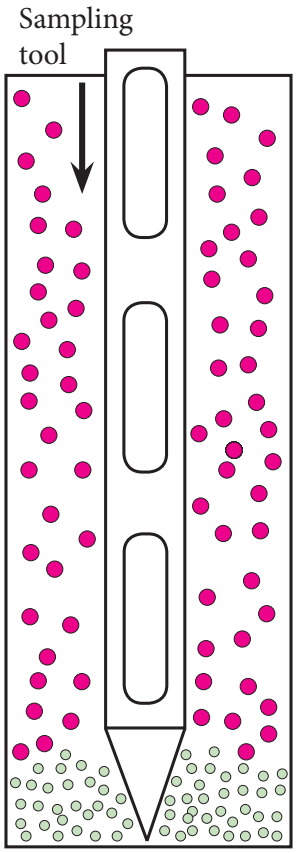


Figure 8. Example where the tool excludes the fine particles that have segregated to the bottom of the container.

purpose) whenever possible (e.g., during loading or unloading, following the adage "sample solids when they are in motion").

An additional design criterion for particulate matter sampling tools is that the design ensures that all particles, regardless of size or location, have an equiprobable chance of being selected. One common problem with coring devices is exclusion of the largest particles. To prevent this, the width of the opening of the coring device must be at least three times the length of the longest particles in the decision unit (or at least two times the length plus 5 mm for fine powders). Another common problem is the exclusion of fine particles (fines) at the bottom of a container. Most tools are inadequate for collection of fines from the bottom of the container due to the design of the point of the sampling tool. An acceptable solution is to convert the material to a one-dimensional decision unit and collect slices.

Sampling tools should possess other characteristics:

- They need to be simple and reliable. Even the best-cared-for tools end up taking a lot of abuse. If the tool fails, the sampling staff will be tempted to improvise, which may lead to collection of an incorrect sample that will not meet the SQC.

- If the tool is to be reused, it must be easy to decontaminate. Tools with small openings, inaccessible openings, or complicated assembly, for example, will be very difficult to decontaminate.

- Tools must be made of a material that does not contaminate the sample or introduce an analyte that causes a matrix interference problem for the laboratory. In addition, the tool material cannot sorb the analyte.

- Tools need to collect increments of equivalent size (mass or volume). If increments are not proportionate, IWE is introduced.

All sampling tools must be well maintained in proper working order.

Automated Sampling Tools. There are many types of automated sampling tools (e.g., Vezin sampling tool, cross-stream sampling tool). Automated sampling tools follow the same requirements for design and use as manual sampling tools. Additional requirements (e.g., speed, cutter angles) are critical considerations in choosing an automated sampling tool. Cleanliness and maintenance is even more critical for automated sampling tools than for manual tools and they require extra attention because contamination and wear is not always readily apparent.

Splitting Tools for Mass Reduction. Three major splitting techniques can be used for mass reduction: rotary splitting, fractional shoveling, and stationary riffle splitting (see Figures 9, 10, and 11). Rotary splitting is the most accurate (because it selects more increments than the other techniques) followed by fractional shoveling and then stationary riffle splitting. Note: Coning and quartering is not discussed because it is a very poor mass reduction method that should never be used.

Splitting Techniques

Rotary Splitter. A rotary splitter is a motorized mechanical device in which the sample is split into hundreds of increments. The device is very simple to operate and relatively inexpensive compared with other laboratory equipment. The rotary splitting technique is not as operator dependent as stationary riffle splitting, but care still needs to be exercised.

Fractional Shoveling. Fractional shoveling is a very simple splitting technique with the following advantages: it does not involve extra equipment (e.g., rifflers); it has minimal cleanup and decontamination requirements; and any number of splits can be generated.

Stationary Riffle Splitter (Jones Riffler). The most common equipment for sample splitting in many laboratories is the stationary riffle splitter. There are two common types of stationary riffle splitters: gated and non-gated. Gated riffle splitters have a trap door between the hopper and the riffles. After the sample is poured into the hopper, the trap door is opened to allow the sample to flow through the rifflers into the receiving pans. A non-gated riffle splitter does not have a trap door, so the material is poured over the riffles and immediately falls through the rifflers into the receiving pans. Of the two types of stationary riffle splitter, it is generally accepted that gated riffle splitters provide more accurate splits. The reason for this is a function of operator error and condition of equipment. Although both types of riffle splitters should theoretically give equivalent results when used properly, the gated riffle splitter is more forgiving of operator error.

All three splitting techniques are applicable for dry granular, powdered, or free-flowing materials such as cereals, grains, nuts, powdered drink mixes, and dry cake mixes. The fractional shoveling method is also applicable for wet materials such as ground meats, thick sauces, and other non-flowing material that will not pass through a splitting device.

Refer to AAFCO's *Guidelines for Preparing Laboratory Samples* manual or to *ISO 6498 Animal feeding stuffs—Guidelines for sample preparation* for specific information on design criteria and proper use of each of these techniques.

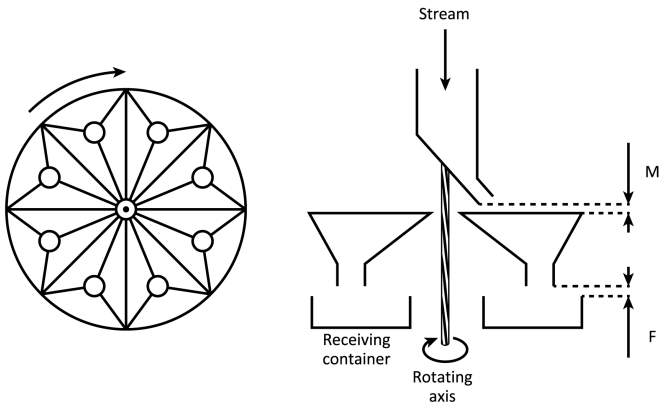


Figure 9. Rotary splitter. Reproduced with permission from Pitard (1993).

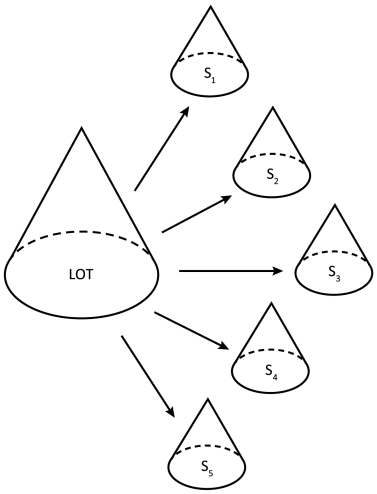


Figure 10. Fractional shoveling. Reproduced with permission of Pitard (1993).

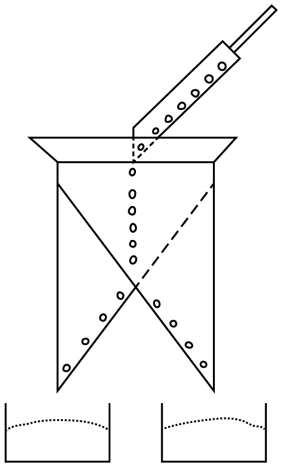


Figure 11. Stationary riffle splitter. Reproduced with permission from Pitard (1993).

EVIDENTIARY (AND ANALYTE) INTEGRITY

Introduction

Evidentiary integrity is the assurance that evidence has not been compromised from sample collection through the generation of analytical results. Evidentiary integrity is the process that ties a test result to a specific decision unit. This is often a legal requirement that is typically initiated when sampling is performed to comply with regulatory purposes. Evidentiary integrity assumes that sample correctness is maintained from primary sample collection through generation of analytical results. Each organization will have evidentiary requirements specific to its needs.

> ***Note:*** Sample correctness must be achieved during all sample mass reduction stages (e.g., primary sample collection, sampling splitting, selection of test portion). Sample correctness must be maintained during all non-mass reduction stages (e.g., particle size reduction, storage). Maintaining sampling correctness during all non-mass reduction stages is addressed in the notion of analyte integrity.

Evidentiary integrity is focused on the "identification and authentication" of evidence. Because analyte integrity is such a major component of evidentiary integrity, it is discussed separately in this document. For the purposes of *GOODSamples*, "analyte integrity" will be the assurance that the material is handled in such a manner that sample correctness is maintained so that it continues to be representative of the decision unit throughout its lifetime; and "evidentiary integrity" will be the assurance that the documentation that tracks the sampling and sample handling process is sufficient for the data to be used as evidence.

Analyte Integrity

To ensure physical, chemical, biological, and/or radiological analyte integrity, communication and agreement are needed among the program staff, sampling staff, and laboratory staff. Considerations for analyte integrity include preservatives, containers, holding times, sampling techniques, and packaging and shipping. To maintain ana-

lyte integrity (control bias to a negligible level), the following should be addressed in a program's specific sampling protocols to minimize bias due to analyte degradation or unintended enrichment:

- Oxidation/reduction: Is oxidation an issue for your analyte? Is so, how can it be prevented?
- Microbial viability: Will the organism grow? Will it die? Is that a problem? How can this be prevented?
- Volatilization: Is the sample or analyte volatile? Is so, how can volatilization be prevented?
- Degradation: Is the sample or analyte subject to degradation? Is so, how can degradation be prevented?
- Contamination/adulteration: Is there potential for accidental contamination of the sample. If so, how can it be prevented during collection, storage, or processing?
- Intentional contamination: How can any tampering be detected? Is the packaging tamper evident?
- Changes in moisture content: Are changes in moisture content critical? If so, how can moisture content be controlled?
- Contamination of the sample during the sampling process: Are techniques required to eliminate potential contamination during the sampling process? Are the tools clean? Is the environment clean? Are the containers clean? Are sampling staff contaminant-free?
- Temperature: Does a change in temperature change the characteristic or analyte of interest?
- Light: Does light affect the analyte or characteristic of interest?
- Packaging and shipping: Is the packaging sufficient to prevent damage to the sample container during shipping and handling?

Preservatives. Sample preservation is action taken to minimize change or loss of analyte(s) of concerns from primary sample collection through disposition. Sample preservation may occur at the collection site or in the laboratory, and should be established

with communication between all parties involved in the sampling process. There are two common types of preservatives: chemical (e.g., pH adjustment) and temperature. It is critical that the preservation method does not interfere with any analysis.

Containers. Containers are used to handle samples at various stages in the sampling and analysis process. The primary sample container should ensure safe arrival at the laboratory; be able to protect the sample from outside contamination, loss, and the environment; and protect handlers from contact with the sample material. Containers used in the laboratory for analytical portions should be also selected with consideration of physical and chemical characteristics. Some important characteristics of containers that may be considered are strength, size, permeability, ability to label, sterility, inertness, and opacity.

Holding Times. Samples may not remain representative of the decision unit indefinitely. When analytes of concern have limited "shelf life," holding times need to be addressed. Preservation techniques retard the chemical and biological changes that inevitably continue after the primary sample is removed from the decision unit. Holding time is analyte specific and may be affected by many things, including preservation, container type, moisture content, stability of the analyte, and sample mass. The laboratory needs to be notified of shipment and the time of collection for samples that are time sensitive.

Sampling Techniques. Sampling may require special techniques, tools, and/or equipment to maintain analyte integrity. Example sampling considerations are

- aseptic collection techniques, if required for microbial testing;
- personal hygiene and personal protective equipment; for example, hand washing, shoe covers, and hair nets; and
- using appropriate tools to avoid contamination or adsorption of analyte; for example, trace-level contaminants.

Evidentiary Integrity

When the data generated from samples collected is used to make decisions, the sampling process and documentation must withstand legal scrutiny. When industry collects samples to comply with regulatory requirements, there is also a critical need for data integrity that will withstand the scrutiny of regulatory bodies.

The purpose of evidentiary integrity is (1) to be able to tie a test result to a specific decision unit; (2) to demonstrate that the sample has not been adulterated or compromised during any step of the process from primary sample collection through generation of the analytical data; and (3) to assure that analyte integrity has been maintained. Evidentiary integrity requires correct sampling. Evidentiary integrity typically consists of (1) documentation (e.g., chain of custody forms, notebooks, worksheets), and (2) procedures (e.g., tamper-evident containers, locked storage, chain of custody).

Documentation. Thorough documentation is a key component to maintaining evidentiary integrity of the entire sampling and analytical process. Such documentation (1) traces the analytical result from the laboratory back to the decision unit; (2) records enough information about the sampling and laboratory preparation process so that it can be recreated to make a determination that the sample did accurately represent the decision unit; and (3) includes information for trace-back investigations, as required. Documentation may occur in various forms, including paper and electronic records.

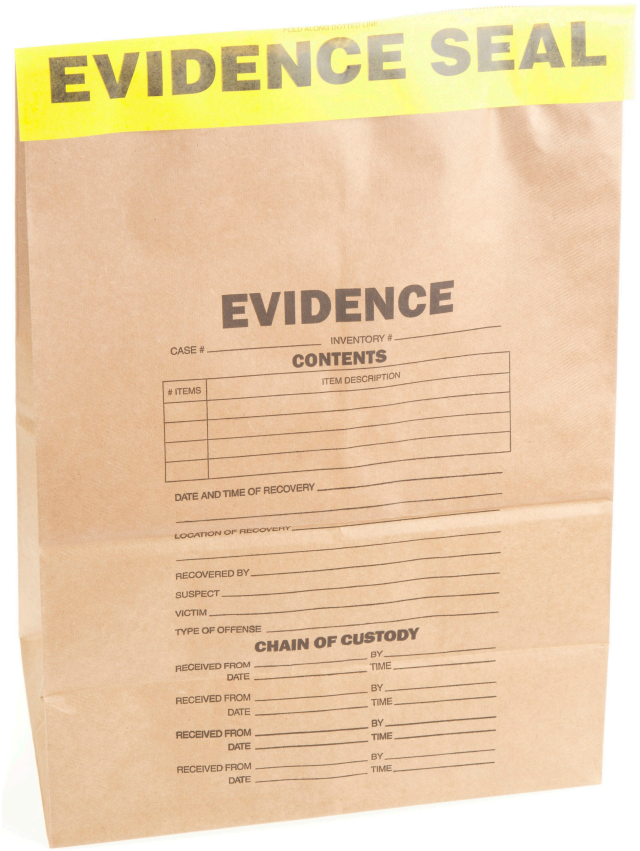
> *"If it isn't written down, it didn't happen."*

The information collected can vary among programs or objectives and needs to be established when developing sampling protocols. Documentation needs to follow the requirements of each organization. Minimum documentation includes the following information: identification of the decision unit, characteristics of the decision unit, date and time of sample collection, identification of all persons involved in sample collection, sample collection technique (e.g., mass, increments, tools), technique for ensuring randomness of increment selection, unique sample identification, and shipping details.

Any deviations from the written protocols must be documented. However, before any deviation or change is implemented, it must first be determined that it would have no deleterious impact on the SQC.

Procedures. Procedures can be established to ensure that the objectives of maintaining evidentiary and analytical integrity are consistently met. Such procedures may include provisions for the use of tamper-evident containers, locked storage space, and documented standard operating procedures for the many associated tasks such as equipment cleaning, sampling protocol, sample preservation, and so on. Procedures include those to ensure analyte integrity, as discussed earlier in this chapter.

Chain of Custody. "Chain of custody is a formally documented continuity of possession, and proof of integrity of evidence collected, which establishes each person having custody/being in possession of the evidence" (McGraw-Hill Concise Dictionary of Modern Medicine, 2002). Chain of custody is specific to the legal requirements of each organization, agency, or program. Chain of custody refers to policies and procedures that must be followed to document the identity and authenticity of the samples and data from collection through reporting of the test results for legal defensibility. The details of chain of custody are beyond the scope of this document; however, it is imperative that all applicable chain of custody requirements are followed.



LABORATORY SAMPLING AND PREPARATION

This section provides guidance to laboratories for preparation of laboratory samples received for analysis. Because the concepts discussed in earlier chapters apply in the laboratory, they will not be repeated. Laboratory personnel should therefore refer to this entire document since the principles of mass reduction in the laboratory are identical to the principles of mass reduction for primary sample collection. This chapter emphasizes common issues or concerns in the laboratory and underscores commonly overlooked sources of error. The AAFCO *Guidelines for Preparing Laboratory Samples* provide greater detail for laboratory sample preparation and, although written for feed materials, the information can be extrapolated to other materials.

The laboratory's primary sampling responsibilities are

(1) to respect the decision unit;

(2) to ensure that analyte integrity is maintained during sample preparation and storage; and

(3) to obtain representative test portion(s) of the laboratory sample received.

Laboratory Sample Quality Criteria

The decision unit is carefully chosen to represent the scale of observation necessary to make a defensible decision. The laboratory must take great care to ensure that this decision unit is respected and maintained during laboratory preparation, sampling, and analysis. Instances where a decision unit might be compromised include inappropriate compositing of laboratory samples during sample preparation to minimize the number of analyses or exclusion/inclusion of a portion(s) of the laboratory sample because of preparation or analytical concerns.

Material Properties

Infinite and Finite Element Materials. Laboratory samples can be received as either finite or infinite element materials. The material properties of the laboratory sample can be altered during sample preparation and analysis. For instance, after particle size reduction, a finite element material generally becomes an infinite element material. An infinite element material tested under a microscope, where individual elements are characterized or measured, would become a finite element material.

Compositional and Distributional Heterogeneity. The compositional heterogeneity of a material does not change unless its particle size is reduced; for instance, it is not changed by stirring, mixing, or splitting. The distributional heterogeneity of a material is altered every time the material is physically manipulated; for example, by shipping, shaking, stirring, pouring, grinding, or splitting.

Mixing Techniques. Mixing is commonly used in the laboratory in an attempt to reduce the distributional heterogeneity of samples; however, the common practice of stirring may increase distributional heterogeneity by promoting segregation due to particle size or density, especially for particulate materials. For liquids and semi-solid materials (e.g., blended canned materials, blended fruits and vegetables), mixing can be an effective technique to reduce distributional heterogeneity.

Theory of Sampling

Whenever mass is reduced (e.g., selecting a test portion or splitting a sample), the principles of minimum mass, minimum number of increments, and sample correctness must be strictly adhered to.

Minimum Mass. The concept of minimum mass to control FSE is as important in the laboratory as it is in collection of the primary sample. This is especially critical in the laboratory where very small test portion masses are taken for analysis. The movement towards miniaturization (smaller and smaller test portions) of analytical methods must consider minimum mass to control FSE. Although collecting random increments is generally considered the greatest challenge in the collection of the primary sample, insufficient mass is generally considered the greatest challenge in the laboratory. The consequences of FSE are frequently ignored at many stages in the laboratory (e.g., generation of analytical samples from laboratory samples, selecting a test portion).

Minimum Number of Increments. The collection of multiple increments is critical in controlling distributional heterogeneity during any mass reduction step. One example of distributional heterogeneity is segregation of elements due to differing size, shape, and density. In the laboratory, collection of random increments is easy due to the small mass (complete accessibility) of most laboratory samples. The common practice of taking a single increment during mass reductions steps is inexcusable and not acceptable for defensible sampling.

Sample Correctness. The principles of sample correctness apply in the laboratory. Laboratory staff should be aware of the effects of dimension and tool shape when performing mass reduction and selecting a test portion, keeping in mind that a one-dimensional or two-dimensional configuration is preferred.

Particle Size Reduction (Comminution) Equipment

Particle size reduction is one of the most important steps in preparation of analytical samples in the laboratory. Particle size reduction is critical to control the FSE. In additional, if comminution produces a uniform shape and size, it will reduce the distributional heterogeneity, especially from segregation. When evaluating comminution equipment, it is critical to ensure (1) that it is of sufficient capacity to process the laboratory sample, (2) that it will reduce the particle size sufficiently to control FSE, and (3) that it will produce a uniform shape and size to control GSE. Another important consideration is the ease of cleaning the equipment between samples.

A wide variety of equipment types are available to achieve particle size reduction. Equipment must achieve the particle size reduction necessary to meet the error requirements established in the SQC. The various types commonly used for feed materials are well described and discussed in AAFCO's *Guidelines for Preparing Laboratory Samples* manual or in *ISO 6498 Animal feeding stuffs—Guidelines for sample preparation*. These commonly include cutting mills, shearing mills, blending mills, and cryogenic mills. There is a great need for better guidance in laboratory comminution equipment, use and performance for food materials. Some types commonly used for food materials are described in the FDA's *Elemental Analysis Manual* and include blenders and homogenizers of various sizes, cryogenic mills, meat choppers, cutting mills, and shearing mills. Because different analytes often have specific, unique preparation needs (e.g., microbiological testing), these needs must be considered in the sampling protocol. Unfortunately, a single type of equipment cannot handle all types of materials and it is imperative that laboratories have adequate equipment to handle the types of materials they will encounter. Listing of equipment in the above publications does not imply that it is suitable for a specific application. Equipment and the method of usage should be validated for a specific application.

Considerations when selecting particle size reduction equipment for a specific application include the following:

- physical and chemical properties of the material;
- initial maximum particle size of the material;
- final desired particle size and the range of permissible particle sizes;
- needed capacity and throughput;
- inertness to analyte of interest;
- complete sample recovery; and
- ease of cleaning, disinfecting, and sterilization.

Evidentiary (and Analyte) Integrity

Analyte Integrity. It is essential to maintain the integrity of analytes of concern in samples throughout their duration in the laboratory. Proper handling and storage may include maintaining proper temperature (refrigeration or freezing), humidity (protection from moisture gain or loss), protection from UV light, and so on. The proper handling and storage conditions will be specific to the materials and analytes of concern. In deciding upon the proper handling, preparation, and storage conditions for each material and analyte combination, laboratories need to consider the effects that composition, matrix interactions, and chemical or enzymatic activity will have on the analyte(s). Handling and storage policies should be established and documented within the laboratory to address these issues.

Maintaining analyte integrity is critical throughout mass reduction and sample preparation stages. Some factors that affect analyte integrity during these stages are (1) heat generation during particle size reduction, (2) adsorption of analyte onto equipment, (3) volatilization of analyte from too-rigorous handling, (4) loss of fine particles due to improper handling, (5) contamination from equipment that contains the analyte of interest (non-inert sampling equipment), and (6) carryover contamination from previously processed samples.

Evidentiary Integrity. No break in evidentiary integrity should occur as primary samples are transferred to the laboratory. The laboratory must have established procedures for maintaining evidentiary integrity from receipt through final sample disposition.

Validation of Laboratory Sample Preparation Protocols

Sample preparation protocols should be validated before implementation. Validation of protocols is important when implementing a new protocol or training personnel involved in the process. Protocol validation is not material or analyte specific and does not negate the need for incorporating sample-specific or process quality control. Successful protocol validation does not imply universal applicability for all materials and analytes; it demonstrates a minimum level of competence and fitness for purpose. Specific examples of performance tests to validate protocols for mass reduction, particle size reduction (comminution), carryover, and mixing can be found in AAFCO's *Guidelines for Preparing Laboratory Samples* and in *ISO 6498 Animal feeding stuffs—Guidelines for sample preparation*.

The terms "grind," "blend," and "mix" can be interpreted very differently. The context in which they are used in *GOODSamples* is as follows:

- **Grind/Grinding:** The act of reducing particle size (comminution) in dry materials using mechanical means (e.g., grinders, shredders, cutters).
- **Blend/Blending:** The act of reducing particle size (comminution) in moist or semi-moist materials using mechanical means (e.g., food processors, blenders).
- **Mix/Mixing:** Stirring, shaking, rolling, and so on, are common practices to attempt to reduce distributional heterogeneity. However, mixing often increases distributional heterogeneity. A better practice to control distributional heterogeneity is to select sufficient increments when performing a mass reduction step.

Performance Tests

Grind/Blend Quality and Recovery. Performance tests for grind/blend quality and recovery estimate the uniformity of the particle size in the ground material and the full recovery of the material. Generally these are accomplished as follows: Weigh a material, pass it through the grinder, pass the ground material through a series of sieves, weigh

the portions retained on the sieves, and calculate recovery; then, calculate the distribution of particle size compared with the desired particle size. Alternatively, an automated particle size analyzer could be used in place of sieves.

Carryover. Performance tests for carryover evaluate the potential for contamination as a result of an inadequate cleaning procedure. Using the procedure, grind/blend a material containing an analyte that is easy to detect (with similar characteristics to the analyte of interest), clean the equipment, and grind/blend a second material containing none of the target analyte. Check or test the second ground material for presence of the target analyte.

Mixing. Performance tests for mixing visually evaluate mixing efficiency. Select at least two materials of varying particle size, density, and different colors. Layer them into a container. Perform normal mixing procedure and make a visual examination for color distribution. Determine the length of time to achieve a visually uniform product. Keep in mind that a visually uniform material does not necessarily imply the absence of distributional heterogeneity.

Reserved Materials

Sometimes a portion of the laboratory sample is reserved in its original condition. There are only two ways to do this properly: (1) The primary sample has sufficient mass so that the FSE of each split is controlled sufficiently to meet the SQC (see Quality Control) or (2) an additional primary sample is collected as the reserve sample. The need for reserve material and the method for collecting the reserve material needs to be specified in the SQC.

QUALITY CONTROL

Quality control is used to assess data quality, monitor process control, and validate methods. Quality control for sampling is essential. Although quality control to estimate error in analytical testing is widely implemented, quality control to estimate error in sampling or sample preparation is rarely implemented. As discussed in the Theory of Sampling chapter, global estimation error (**GEE**) is the square root of the sum of squares of all contributing errors, including sampling, sample preparation, and analytical errors, among others. Ignoring the contribution of sampling error to the GEE is a serious oversight because sampling error often exceeds analytical error. Use of the analytical error as the only estimate of GEE is an incorrect practice that must be remedied. The implementation of quality control throughout the entire process from primary sampling through testing is necessary to determine GEE.

Application of Quality Control

Quality control can be implemented (1) to determine the GEE for individual samples; (2) to determine if a process is in control; or (3) to validate a method or protocol. These are three different concepts. Quality control for individual samples is used to estimate the error in the test result for a specific decision unit. Quality control for a process determines if the process is in control, with no specific information about individual samples. Process quality control has specific requirements, which will not be discussed in *GOOD-Samples*. Quality control to determine the GEE will be addressed here in detail.

Quality control can be process-wide (covering all steps from primary sampling through testing) when a single measurement of GEE is desired. Alternatively, it can be implemented at any mass reduction stage in the process when it is desired to estimate the error contributions from that specific stage in the process.

Estimation of Error. Every step in the sampling process (e.g., collection, splitting, containerizing, mass reduction, preparation) introduces error. This can include errors of bias (due to incorrect sampling tools and equipment, or contamination), errors of imprecision (due to insufficient mass and increments), or gross errors (due to sample swapping, spillage, and so on). Errors of precision can be monitored through replication. Bias

errors and gross errors are very difficult to measure and every effort should be made to avoid their introduction or minimize them to the point that they are negligible.

The type and frequency of quality control events may vary among organizations. Each organization needs to consider what sampling, sampling preparation, and analytical quality control to perform, as well as how to interpret quality control events. Each organization also needs to determine a frequency for specific quality control events. The relative risk of incorrect decisions, risk of contamination, heterogeneity of the material, required confidence, and so on, are considerations that will determine the type and frequency of quality control events.

When performing quality control to estimate error, it is critical that all errors be accounted for. Errors that are relatively small compared with other errors have little effect on the total error. Because sampling-related errors are typically greater than testing-related errors, it is imperative that sampling quality control be incorporated in the estimation of error in the final test result.

Bias. Bias comes from two sources: the collection of the sample (e.g., use of incorrect tools) and from factors external to the sample collection (e.g., contamination, loss of analyte integrity). Although it may be possible to detect bias, the quantification of bias is difficult, if not impossible. Therefore, elimination of bias is critical in sampling and sample preparation. There are no quality control events to determine bias from incorrect sample tool design or use. Therefore, it is critical that every organization procures appropriately designed tools and prioritizes training on proper selection, use, and maintenance of tools. Sampling biases are inconsistent and therefore manifest as imprecision.

Contamination error is one form of bias that may be detected. The presence or absence of contamination in a quality control blank does not prove or disprove contamination of a sample. Rather, it provides evidence that contamination has occurred.

Contamination Quality Control. Blanks are used to check for contamination. The contamination can be from the containers or environment in which the samples are collected, vehicles used for transportation, and (or) carryover from tools and equipment. Blanks can be created in various ways and, although there are some common ways blanks are created, there is no set rule; each organization is free to establish their own system as long as they can demonstrate that they can detect contamination. Blanks are submitted to the laboratory with the primary samples.

The frequency of inclusion of blanks is determined by the probability of contamination and the required confidence in detecting the presence of contamination. For samples less susceptible to contamination, contamination checks may be performed less frequently; for samples more susceptible to contamination, contamination checks may be performed more frequently. For high-concentration analytes (e.g., 10%), contamination checks may be performed infrequently, if at all; for low-concentration analytes [e.g., mg/kg (ppm) or µg/kg (ppb)], contamination checks may be performed more frequently. If the consequences of non-compliance (financial, legal, or regulatory) are severe, quality control contamination checks should be performed more frequently.

Quality control checks for contamination from containers depend on

- **The cleanliness of the containers:** If the supplier tests the containers and provides certification, contamination checks may be performed infrequently, if at all. If the containers are new, but not certified, perform checks more frequently. If containers are re-used, perform checks yet more frequently. Note: Use of recycled containers is discouraged.

- **The integrity of the container:** If there is likelihood a target analyte can leak or permeate into or out of the container, then quality control checks for contamination should be performed. Historically, for many analytes this has not been considered an issue. However, for new or emerging analytes of lower concentrations, this must be considered.

Quality control checks for environmental contamination depend on

- **The environmental conditions:** These may include atmospheric contamination or dirty surfaces. If the sampling environment is considered "dirty," contamination quality control would be performed more frequently than when sampling in an environment that is considered "clean."

- **The sampling staff:** The relative cleanliness of the sampling staff and their practices in maintaining a clean work environment (e.g., failure to change gloves immediately before sample collection, or placing clean tools and containers on dirty surfaces) will affect sample quality.

Quality control checks for carryover contamination from tools and equipment depend on

- **The concentration of the analyte of concern.** When collecting or preparing multiple samples, when possible work from low concentrations of the analyte(s) of concern to progressively higher concentrations.

- **The complexity of tools or equipment in terms of trapping residual material.** Some tools can be opened, cleaned and visually inspected. Other (less desirable) tools and equipment cannot be opened, easily cleaned, or visually inspected for cleanliness. Whenever possible, choose tools and equipment that can be visually inspected.

- **The chemical and physical characteristics of the analyte of concern.** If the analyte(s) of concern has a propensity to stick or sorb onto the equipment surface, blanks are more necessary.

Blanks are generally collected by swabbing or rinsing all surface areas that come in contact with the sample or they can be created by processing an inert material with the equipment and testing it for the contaminant. Tool or equipment blanks are collected after the tool or equipment is considered clean for use.

Imprecision. Fundamental sampling error (FSE) and grouping and segregation error (GSE) are the primary mass reduction imprecision errors and are estimated through replication. Replication can be initiated at any stage in the sampling process. Replication initiated at any point reflects the error in that stage plus the error in all subsequent stages. Therefore, replication at multiple stages can be used to isolate error at a specific stage. The replicates must be collected using exactly the same protocol. Therefore, when replicating a primary sample or test portion, the increments for each replicate must be collected from random locations (not co-located with or split from other increments).

Replicated primary samples can be used to generate multiple test portions that provide an estimate of the GEE. Therefore, replication at the primary sampling stage should be implemented whenever possible. Replicating test portions from the same analytical sample provides no information about imprecision associated with all the preceding stages.

Imprecision error is typically expressed as a relative standard deviation (**RSD**) or percent coefficient of variation (**%CV**). A minimum of three replicates is recommended for all imprecision estimates; however, more replication yields better estimates.

Quality Control Replicates

Whenever mass reduction occurs, quality control can be implemented to estimate the imprecision of the mass reduction stage. Each project will have a unique series of sample preparation and mass reduction stages. The following is a generic illustration for

triplicates and does not imply that these should be the sample preparation/mass reduction stages for any given project. This process can be extended for any number of replicates and any number of sample preparation and mass reduction stages (see Figure 12).

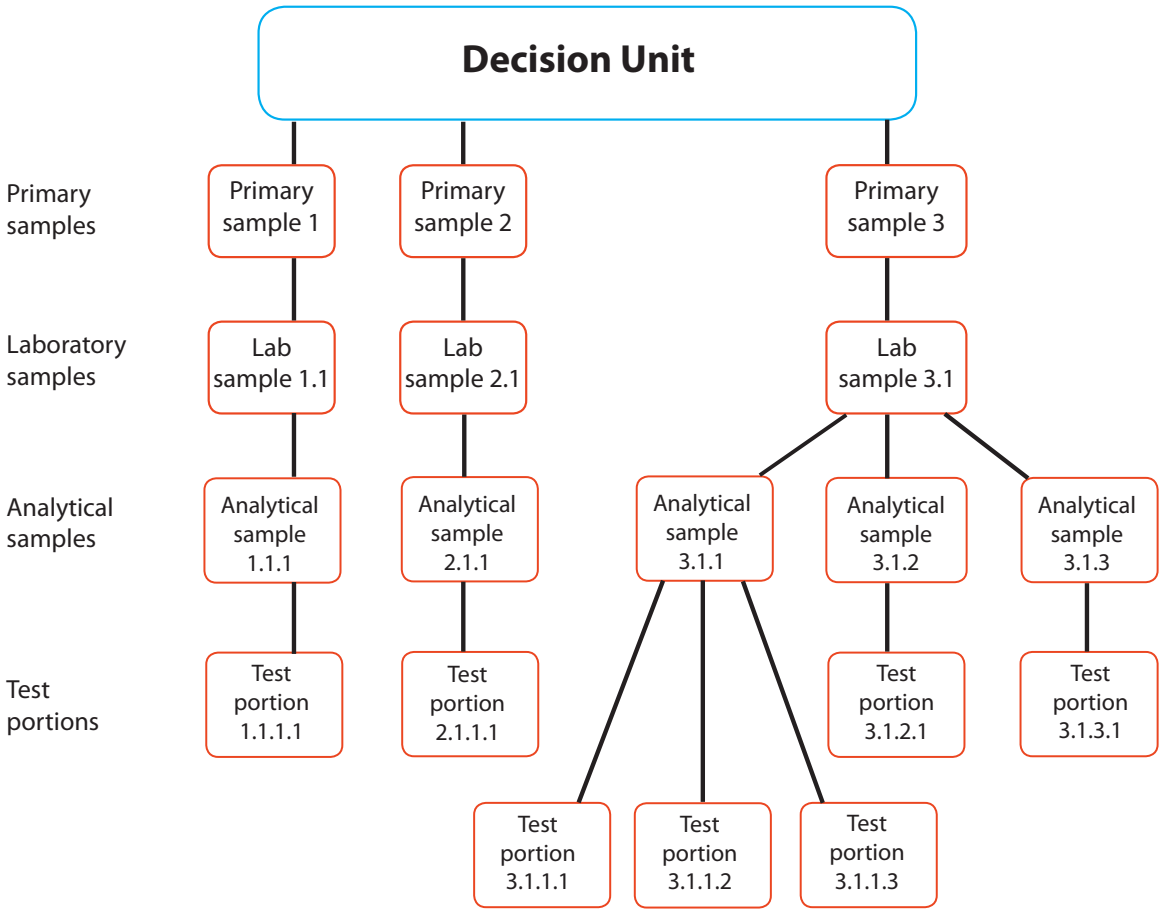


Figure 12. Levels of replication using triplicates.

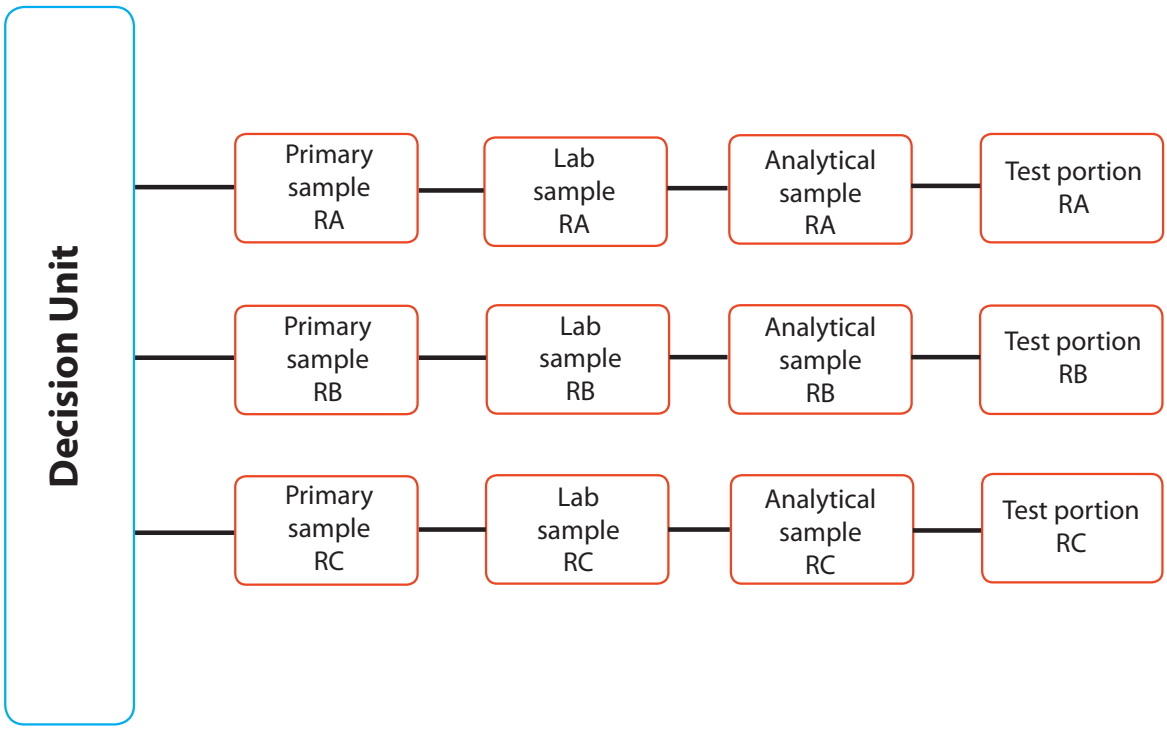


Figure 13. RA, RB, and RC are replicate primary samples. Data resulting from test portions RA, RB, and RC can be used to estimate the GEE associated with the repeatability of the entire process, from sampling through analysis.

- **Primary Sample Replication:** Replicated primary samples can be used to generate multiple test portions that provide an estimate of the GEE. Therefore, replication at the primary sampling stage should be implemented whenever possible.

- **Analytical Sample Replication:** Replicated analytical samples can be used to generate multiple test portions that provide an estimation of the imprecision associated with selection of the analytical sample, the selection of the test portion, and the test. Replicating analytical samples provides no information about imprecision associated with preceding stages.

- **Test Portion Replication:** Replicating test portions from the same analytical sample provides an estimation of the imprecision associated with selection of the test portion and the test. Replicating test portions provides no information about imprecision associated with all the preceding stages.

Split Sample. "Split sample" is a term that is used in many contexts. Splitting is a mass reduction technique and all rules of mass reduction apply. Splitting is a fundamentally different process than replication. The goal of replication is to estimate error; the goal of a splitting is to produce identical portions of the original material. Splits are done to make post-split comparisons (e.g., compare analytical methods, compare comminution methods, compare labs, or compare analysts). The comparisons become meaningless if the splits are different because the splitting error cannot be isolated from the error under study. Any time a split is generated, it must have sufficient mass and number of increments to control FSE and GSE. It is also imperative that the principles of sample correctness be adhered to during all splitting processes. Figures 13 and 14 illustrate the difference between three replicate primary samples and three splits of primary sample from a decision unit.

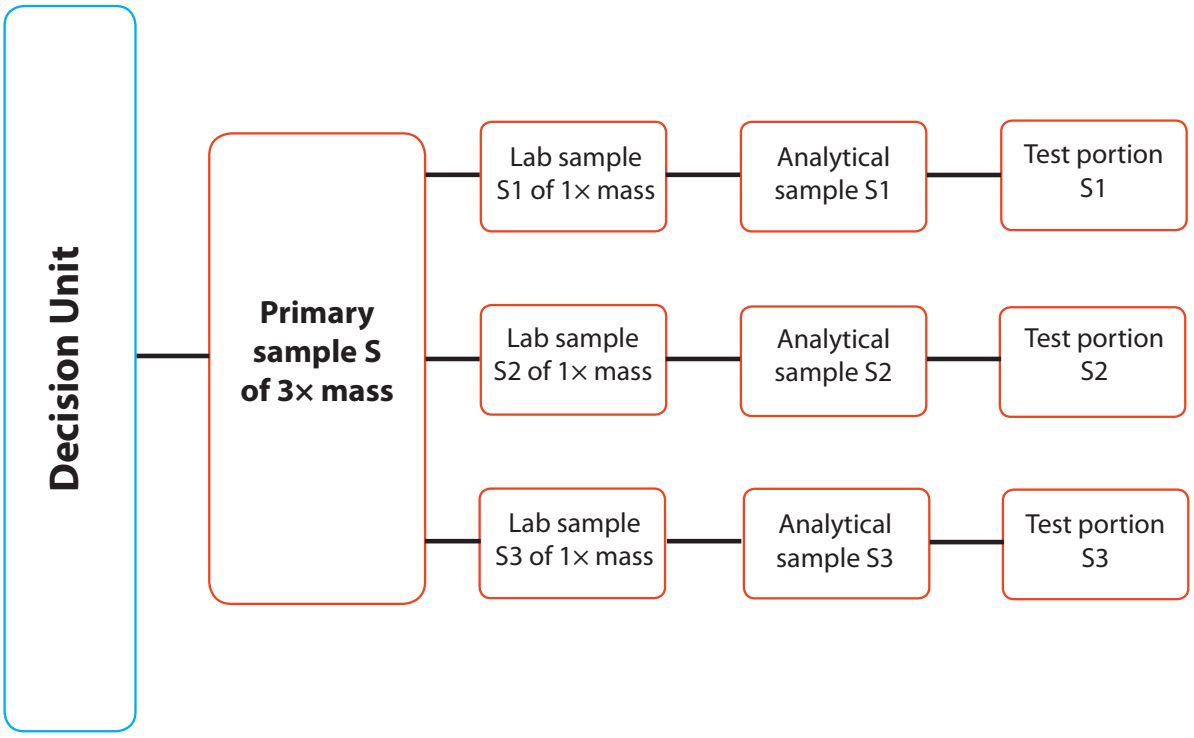


Figure 14. Primary sample S is divided to produce split laboratory samples S1, S2, and S3. Generally, the split should be performed at the stage preceding the comparison stage. Data resulting from test portions S1, S2, and S3 can be used to compare three different procedures or methods of preparing the analytical sample. To generate three splits, the primary sample must be at least 3× the mass so that each laboratory sample is of sufficient mass to represent the decision unit.

INFERENCE

Inference is the process of estimating (or inferring) a concentration or characteristic about a decision unit based on a sample(s) collected from that decision unit. There are two forms of inference significant in food and feed sampling.

One form of inference is the estimation of an average analyte concentration of a decision unit from a sample of that decision unit. The quality of this inference is a function of the amount of error introduced in the sampling and analytical process. The more representative the sample (the lower the error introduced), the higher the quality of the inference (or the greater the confidence). The lower the error, the closer the test result is to the true analyte concentration in the decision unit.

A second form of inference is the estimation of the percentage (proportion) of decision units that have some specific concentration or characteristic based on sampling multiple decision units. This form of inference is common when there are more decision units that can be sampled but inference is desired to all the individual decision units, sampled or not. For this form of inference, there must first be an inference to sampled decision units. The greater the number of decision units sampled, the higher the quality of the inference to the unsampled decision units.

Inference to Decision Unit

Inference to a decision unit (direct, probabilistic, or statistical) is based on a sample(s) from the decision unit. Inference occurs at every mass reduction stage from the analytical result to the decision unit in the sampling to testing pathway (e.g., inference from the analytical result to the test portion, inference from the test portion to analytical sample, inference from analytical sample to primary sample, inference from primary sample to decision unit). The number of inference steps is unique for each sampling project.

Direct Inference. If the entire decision unit is collected as the primary sample, it would be a "perfect" sample that contains no error (assuming analyte integrity is maintained). If the entire primary sample is analyzed in its entirety, this yields the best possible inference and is termed *direct inference*.

Probabilistic or Statistical Inference. If the entire decision unit is collected as the primary sample but it is not analyzed in its entirety, the test portion selected from the laboratory or analytical sample needs to be representative to make inference back to the decision unit. If the entire decision unit is not collected as the primary sample, the portion selected from the decision unit needs to be representative to

make inference back to the decision unit. These types of inferences are *probabilistic inference* and *statistical inference*. The amount of error introduced will affect the quality of the inference. The tolerable GEE should be specified in the SQC and sample correctness must be ensured and maintained or any inference is invalid.

Inference to Unsampled Decision Units

If there are more decision units than can be sampled, inference can be used to estimate (infer) the proportion of decision units that have a specific concentration or characteristic (Figure 15). This type of sampling is sometimes called acceptance sampling. A special application of acceptance sampling can be used to determine if an acceptably small proportion of the decision units have a certain concentration or characteristic. This type of sampling is commonly used in cases where any detection of an analyte is unacceptable. Note: No sampling protocol can determine with absolute certainty that all decision units meet specific criteria unless every decision unit is sampled.

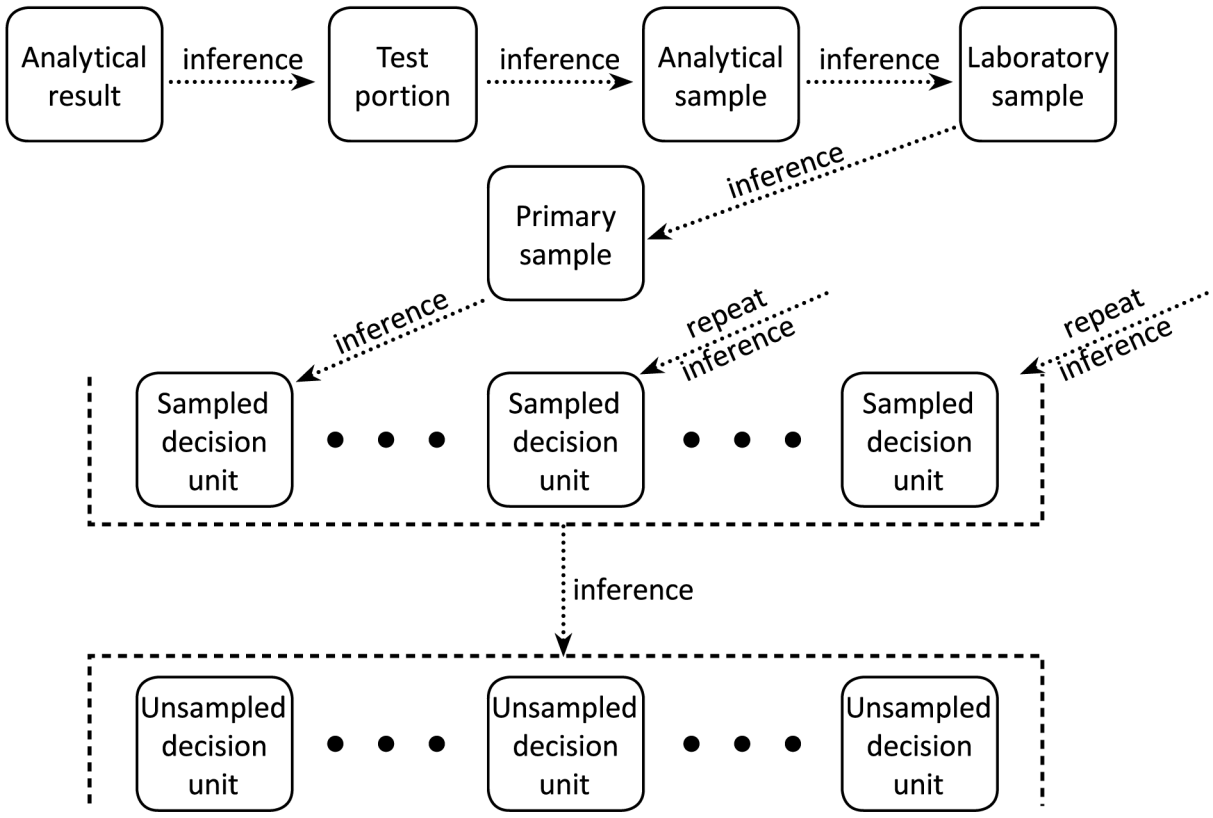


Figure 15. Examples of forms and stages of probabilistic inference to a decision unit and to unsampled decision units. (Reproduced with permission from J. AOAC Int. 98(2), 288–294. doi: http://dx.doi.org/10.5740/jaoacint.14-292.)

Limitations of Inference

As only a portion of the decision unit (for inference to a decision unit) or a portion of the decision units (for inference to unsampled decision units) is selected for inference, the possibility always exists that an incorrect inference will be made. More representative samples (within a decision unit) and more decision units sampled (from all decision units) will reduce the rate of incorrect inference, but it will not eliminate it. For example, if 500 decision units are sampled (out of a greater number) and none of the 500 decision units shows the presence of bacteria, it does not imply that no bacteria are present in any of the decision units; it only implies that the presence of bacteria is a relatively rare event, if present at all.

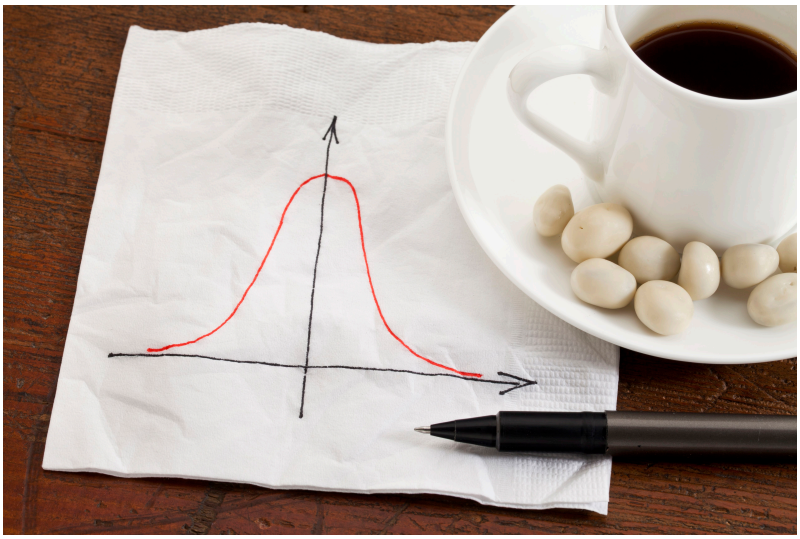
DATA QUALITY ASSESSMENT

Data quality assessment is a procedure used to evaluate if the data is of sufficient quality to meet the requirements set forth in the sampling quality criteria (**SQC**). Determining that the data meet the SQC involves review of documentation and determining global estimation error (**GEE**). To determine the GEE, quality control data must be evaluated. It is imperative that sample correctness be evaluated as part of assessment. If sample correctness is not ensured and maintained, valid inference and defensible decisions cannot be made.

Review of Documentation

Data quality assessment generally begins with a review of documentation, including documentation associated with SQC, protocols, and evidentiary integrity. This would constitute a review of all forms of documentation (e.g., field books or sheets, chain of custody forms, bench sheets) to determine if all tasks were completed appropriately and documented. Any undocumented task must be considered an omission.

During assessment, deviations or omissions from written protocols must be evaluated with respect to their impact upon the analytical results, GEE, and the ability to make inference. Potential deviations include the inability to collect increments at random due to inaccessibility of the entire decision unit, potential contamination from outside sources, unavailability of the desired decision unit(s), failure to use the correct sampling tool, breakage of containers, spillage of samples, or deviation in transportation or storage conditions from those specified.



Evaluation of Quality Control and Global Estimation Error

Evaluation of Blanks. The presence of contamination in blanks indicates the potential for contamination in samples. The absence of contamination in blanks does not prove that there is no contamination in samples, but supports the assessment that the probability of contamination risk is low.

If contamination exists in the blank samples, a determination must be made whether the contamination has a significant impact on the analytical results. Incidental contamination may be insignificant for high-level concentrations. For example, a parts-per-billion (ppb) level contaminant of a target analyte present at parts per thousand may not have a significant impact. If contamination is a common observation, even at insignificant levels, the sampling and sample handling processes should be evaluated.

Evaluation of Replicates and Global Estimation Error. Replicates are used to estimate the imprecision associated with various mass reduction stages. Replication at the primary sampling stage provides an estimate of the GEE. There are countless possibilities for SQC and quality control and therefore specifics on data quality assessment are complicated. Some general considerations and cautions that are universally applicable follow.

How Does the Actual GEE Compare to the Specification Limit or Concentration of Concern? During the development of the SQC, the actual concentration of the analyte of interest is unknown and the tolerable error is estimated with some assumptions about the true concentration. The error that can be tolerated is only an educated estimate used to develop the sampling protocol. During the SQC process, when the actual concentration is known, it will become evident if the error specified in the SQC is acceptable. A defensible decision is dependent upon the proximity of the actual concentration to the specification limit or action level and the actual GEE (see Figure 16).

Exercising Caution in Data Assessment: Is the Actual GEE Greater than 35% RSD? There is an upper limit to the magnitude of the GEE. Once the GEE exceeds 35% RSD, the results begin to depart from a Gaussian distribution. When the data deviate from a Gaussian distribution, it becomes increasingly difficult to make inference. Once the overall GEE exceeds 100% (the error is greater than the number itself), it is virtually impossible to make

meaningful inference. There is a gradual decrease in the ability to make inference from a GEE RSD of 35 to 100%. As the GEE increases from 35 to 100%, the concentration needs to be further and further away from the concentration of concern for any possible inference. There are no specific formulas or calculations to determine data acceptability for RSD > 35%, but professional judgment needs to be applied. Every organization needs to have policies for data handling when the GEE exceeds 35%. An RSD > 35% is a red flag that resulting decisions may be incorrect; it is a dangerous situation for food and feed safety initiatives and should be remedied.

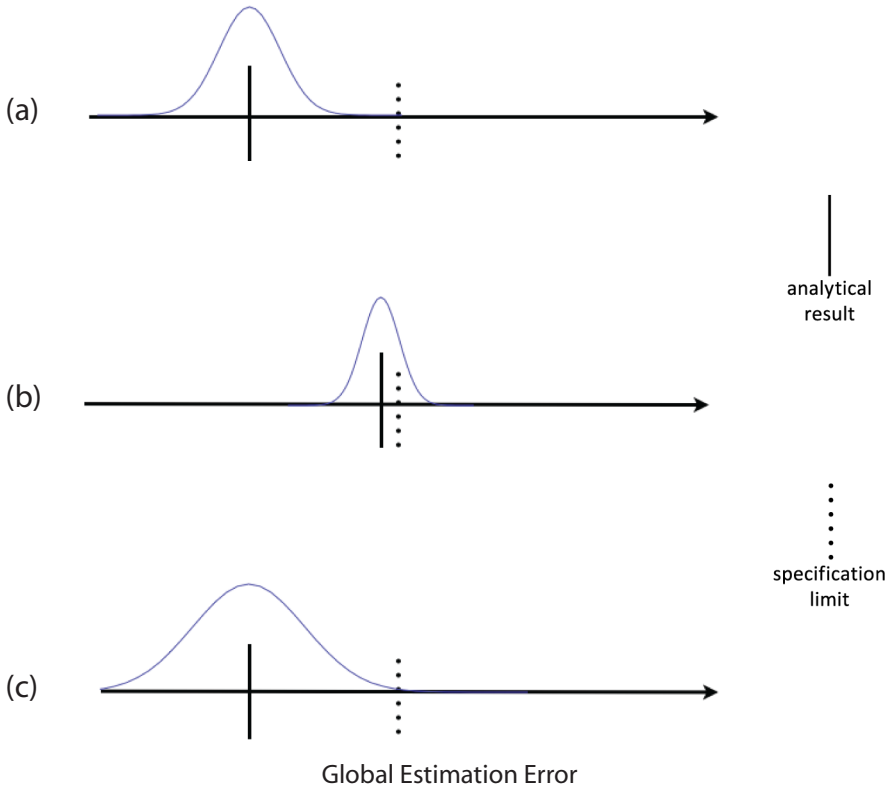


Figure 16. Relationship of result and GEE to the specification limit: (a) acceptable error; (b) unacceptable error; (c) acceptable error. The curve represents the magnitude of imprecision error.

In Figure 16a, the measured concentration and GEE are sufficiently removed from the specification limit and a defensible decision can be made. In Figure 16b, the measured concentration and GEE are insufficiently removed from the specification limit. No defensible decision can be made even though the GEE is much less than in example (a). In Figure 16c, the measured concentration and GEE are sufficiently removed from the specification limit and a defensible decision can be made, even though actual GEE is much larger than in example (a).

Outlier Treatments. As part of data assessment, there may be a temptation to remove data that make interpretation or assessment difficult. Often these data are termed "outliers" and may be inappropriately rejected via statistical outlier tests or the data are replaced by retesting, hoping that the undesirable result can be dismissed or averaged out. However, such data indicate excessive error somewhere in the process and may indeed be the most important result. It is important to determine the source of the error before any actions are taken. It is scientifically incorrect to retest only data that are deemed undesirable or unexpected. If the source of error from an outlier occurrence can be identified, the source of error must be eliminated before any retesting.

Acceptance Sampling. The details of acceptance sampling are widely available in the literature and are not discussed in this document. However, one of the underlying assumptions in acceptance sampling is that there is no error in the estimate of the characteristic or analyte concentration in the decision unit. Because error does exist in the estimate, this error must be measured and the effect of the error incorporated into any acceptance sampling protocol.

When implementing acceptance sampling protocols, the goal is to determine the percentage of decision units that do or do not have a certain characteristic or concentration. Acceptance sampling cannot determine specifically which decision units do or do not possess the characteristic or concentration. It is therefore inappropriate to selectively remove the sampled decision units where the characteristic or concentration is above a threshold and conclude that the remaining material is acceptable. It is never feasible to say that 0% of decisions units (or 100% of the decisions units) are acceptable unless every decision unit is tested.

Example: A field of lettuce contains 10,000 individual heads of lettuce where each head of lettuce is a decision unit. A total of 100 individual heads of lettuce are sampled and tested, resulting in 5 specific heads of lettuce that possess an undesirable characteristic or concentration. The best estimate is that 5% of the 10,000 heads of lettuce (500 heads) have the undesirable characteristic or concentration. Therefore, removal of just the 5 undesirable heads of lettuce (and/or those in close proximity of the 5 heads) does not imply that the remaining 9,995 heads of lettuce are acceptable. The best estimate would be that there are still 495 heads of lettuce in the field with the undesirable characteristic/concentration.

RESOURCES

Primary Resources

Gy, P. (1998). Sampling for Analytical Purposes. New York: John Wiley & Sons.

Pitard, F. F. (1993). Pierre Gy's Sampling Theory and Sampling Practice. 2nd ed. Boca Raton, FL: CRC Press.

Additional Resources

Esbensen, K. H. (2013). DS 3077, Representative Sampling–Horizontal Standard. Danish Standards. www.ds.dk.

Esbensen, K. H., Paoletti, C., & Thiex, N, *Representative Sampling for Food and Feed Materials: A Critical Need for Food/Feed Safety* (2015) *J. AOAC Int.* **98**, 229–251. http://dx.doi.org/10.5740/jaoacint.SGE_Esbensen_intro.

Kuiper, H., and Paoletti, C. (2015). Food and Feed Safety Assessment: The Importance of Proper Sampling. *J. AOAC Int.* **98**, 252–258. http://dx.doi.org/10.5740/jaoacint.15-007.

Thiex, N., Esbensen, K. H., & Paoletti, C. (2015). Towards a Unified Sampling Terminology: Clarifying Misperceptions. *J. AOAC Int.* **98**, 259–263. http://dx.doi.org/10.5740/jaoacint.14-290.

Ramsey, C., & Wagner, C. (2015). A Systematic Approach to Representative Sampling. *J. AOAC Int.* **98**, 264. http://dx.doi.org/10.5740/jaoacint.14-239.

Ramsey, C., & Wagner, C. (2015). Sample Quality Criteria. *J. AOAC Int.* **98**, 265–268. http://dx.doi.org/10.5740/jaoacint.14-247.

Esbensen, K. H. (2015). Materials Properties: Heterogeneity and Appropriate Sampling Modes. *J. AOAC Int.* **98**, 269–274. http://dx.doi.org/10.5740/jaoacint.14-234.

Wagner, C., and Esbensen, K. H. (2015). Theory of Sampling—Four Critical Success Factors Before Analysis. *J. AOAC Int.* **98**, 275–281. http://dx.doi.org/10.5740/jaoacint.14-236.

Esbensen, K. H., and Ramsey, C. (2015). Quality Control of Sampling Processes—A First Foray; From Field to Test Portion. *J. AOAC Int.* **98**, 282–287. http://dx.doi.org/10.5740/jaoacint.14-288.

Ramsey, C. (2015). Considerations for Inference to Decision Units. *J. AOAC Int*. **98**, 288–294. http://dx.doi.org/10.5740/jaoacint.14-292.

Paoletti, C., and Esbensen, K.H. (2015). Distributional Assumptions in Agricultural Commodities— Development of Fit-for-Decision Sampling Protocols. *J. AOAC Int*. **98**, 295–300. http://dx.doi.org/10.5740/jaoacint.14-250.

Wagner, C. (2015). Critical Practicalities in Sampling For Mycotoxins in Feed. *J. AOAC Int*. **98**, 301–308. http://dx.doi.org/10.5740/jaoacint.14-235.

Ramsey, C. (2015). Considerations for Sampling Contaminants in Agricultural Soils. *J. AOAC Int*. **98**, 309–315. http://dx.doi.org/10.5740/jaoacint.14-268.

Ramsey, C. (2015). Considerations for Sampling of Water. *J. AOAC Int*. **98**, 316–320. http://dx.doi.org/10.5740/jaoacint.14-251.

Lehotay, S. J., & Cook, J. M. (2015). Sampling and Sample Processing in Pesticide Residue Analysis. *J. Agric. Food Chem*. **63**(18), 4395–4404. http://dx.doi.org/10.1021/JF5056985.

Ramsey, Charles A. (2015). The Decision Unit – A Lot with Objectives. *TOS Forum* **5**. http://dx.doi.org/10.1255/tosf.75.

Ramsey, Charles A. (2015). The Role of Inference in Food Safety. *TOS Forum* **5**. http://dx.doi.org/10.1255/tosf.77.

Ramsey, Charles, and Thiex, Nancy. 2014. Improved Food and Feed Safety through Systematic Planning and the Theory of Sampling (TOS): An Introduction to "GOOD-Samples". *TOS Forum* **2**.

ISO (2012). ISO 6498:2012—Animal Feeding Stuffs–Guidelines for Sample Preparation. Geneva, Switzerland: International Organization for Standardization.

Thiex, N., Novotny, L., Ramsey, C., Latimer, G., Torma, L., & Beine, R. (2000). Guidelines for Preparing Laboratory Samples. Oxford, IN: Association of American Feed Control Officials.

Feed Inspector's Manual (2014). AAFCO Inspection and Sampling Committee, Association of American Feed Control Officials, Samplers Manual. 5th ed. http://www.aafco.org/.

Petersen, L., Dahl, C., & Esbensen, K. H. (2004). Representative mass reduction in sampling—A critical survey of techniques and hardware. In Special Issue: 50 years of Pierre Gy's Theory of Sampling. Proceedings: First World Conference on Sampling and Blending (WCSB1). Esbensen and Minkkinen (Eds). *Chemometrics and Intelligent Laboratory Systems* 74(1), 95–114.

Esbensen, K. H., & Wagner, C. (2014). Theory of Sampling (TOS) vs. Measurement Uncertainty (MU)—A call for integration. *Trends Anal. Chem.* **57**, 93–106. http://dx.doi.org/10.1016/j.trac.2014.02.007.


GOOD
Samples