CODEX ALIMENTARIUS
INTERNATIONAL FOOD STANDARDS

E-mail: codex@fao.org - www.codexalimentarius.org

GUIDELINES FOR THE DESIGN AND IMPLEMENTATION OF NATIONAL REGULATORY FOOD SAFETY ASSURANCE PROGRAMME ASSOCIATED WITH THE USE OF VETERINARY DRUGS IN FOOD PRODUCING ANIMALS

CAC/GL 71-2009

Adopted 2009. Revision 2012, 2014.

Table of Contents

1. INTRODUCTION
2. SCOPE
3. GENERAL PRINCIPLES
4. APPROACH BASED ON RISK
5. DEFINITIONS (FOR THE PURPOSE OF THESE GUIDELINES)
6. REGULATORY FRAMEWORK
6.1 Roles
6.2 Approval by competent authority
6.3 Information on veterinary drugs
6.4 Sale and use
6.5 Responsibilities of business operators (best practice guidance)
7. VERIFICATION PROGRAMMES
7.1 Purpose
7.2 General design principles
7.3 System and targeted verification programme design
7.4 Risk profiling
8. CHOICE OF VERIFICATION PROGRAMME
8.1 System verification programmes
8.2 Risk targeted verification programmes
8.3 Surveys
8.4 Review
9. SAMPLE TAKING
9.1 General principles
9.2 Traceability/product tracing
10. STATISTICAL CONSIDERATIONS
10.1 General
10.2 Retention of consignments during laboratory analysis
10.3 Result interpretation
10.4 Port of entry testing programmes (specific requirements)
11. REGULATORY ACTION
11.1 Investigation of non-compliances
11.2 Measures in case of non-compliance: Conduct
11.3 Measures in case of non-compliance: Product
11.4 Corrective action in case of non-compliance
12. INTERACTION BETWEEN THE CONTROL PROGRAMMES OF TWO COMPETENT AUTHORITIES

Analytical methods for residue control

General considerations on analytical methods for residue control
13. INTRODUCTION
14. INTEGRATING ANALYTICAL METHODS FOR RESIDUE CONTROL
15. CONSIDERATION FOR SELECTION AND VALIDATION OF ANALYTICAL METHODS
15.1 Identification of Methods Requirements
15.2 Implementing other Codex Alimentarius Commission Guidelines
15.3 Method Validation and Fitness for Purpose
15.4 Single Laboratory Validation – The Criteria Approach

Attributes of analytical methods for residues of veterinary drugs in foods
16. INTRODUCTION

17. METHOD DEVELOPMENT CONSIDERATION

18. ANALYTICAL PERFORMANCE CHARACTERISTICS

18.1 Performance Characteristics of Screening Methods

18.2 Performance Characteristics for Quantitative Methods

18.3 Performance Characteristics for Confirmatory Methods

18.4 General Performance Characteristics for Methods for Use in a Regulatory Control Programme

19. METHOD DEVELOPMENT AND VALIDATION CONSIDERATIONS FOR RESIDUE CONTROL METHODS

19.1 Selection of Appropriate Test Material for Validation

19.2 Measurement Uncertainty

19.3 Use of Internal Standards

19.4 Environmental Considerations

19.5 Choice of Validation Model

19.6 Quality Control and Quality Assurance

Appendix A – Sampling strategies

A1. NON-BIASED SAMPLING

A1.1 Purpose

A1.2 Statistical considerations on sampling population size

A1.3 Sampling Confidence reporting

A2. DIRECTED OR TARGETED SAMPLING

A2.1 Purpose

Appendix B – Sampling of commodities

B1. SCOPE

B2. DEFINITIONS

B3. SAMPLING PROCEDURES

B4. SPECIFIC SAMPLE PREPARATION INSTRUCTIONS FOR HONEY

B5. STATISTICAL CONCERNS

B5.1. Stratified random sampling

B5.2 Systematic sampling

B5.3 Biased or estimated worst case sampling

B6. PREPARATION OF LABORATORY SAMPLES

B7. SHIPMENT OF LABORATORY SAMPLES

B8. RESULT INTERPRETATION IN THE LABORATORY

B9. SAMPLING RECORDS

Guidance on sample type and quantity for different commodities

Table A: Meat and poultry products

Table B: Milk, eggs and dairy products

Table C: Aquaculture products

Appendix C – Performance Characteristics for Multi-Residue Methods (MRMs) for Veterinary Drugs

C1. SCOPE

C2. DEFINITIONS

C3. SUMMARY OF PERFORMANCE PARAMETERS TO BE CHARACTERISED AND DEFINED FOR MULTI-RESIDUE ANALYTICAL METHODS

C4. PERFORMANCE CHARACTERISTICS FOR MRMS

C5. PERFORMANCE CHARACTERISTICS OF MRMS FOR SCREENING ANALYSIS

C6. PERFORMANCE CHARACTERISTICS OF MRMS FOR QUANTITATIVE ANALYSIS

C7. PERFORMANCE CHARACTERISTICS FOR MRMS FOR CONFIRMATORY METHODS

C8. VALIDATION OF THE FULLY CHARACTERIZED MRMS

1. INTRODUCTION

Modern food production systems should be designed and managed to ensure that the exposure of food producing animals to veterinary drugs does not pose a risk to human health.

The commercial entities involved in the production and marketing of food have the primary responsibility for ensuring food safety. The role of competent authorities is to control the use of veterinary drugs and to verify that appropriate practices are being applied and effective measures are in place within the veterinary drug distribution and food production systems to provide effective protection for consumer health and ensure fair practice in the food trade, consistent with the goals of Codex Alimentarius. All parties also have a responsibility to provide consumers information and education to facilitate sound choice of food products of animal origin.

The application of a programme based on risk to all food types should provide the controls and verification consistent with the risk that the food type may pose to consumers. The application of an approach based on risk across all food groups and hazard classes should allow a more focussed application of resources to those areas which are most likely to generate real human health protection gains.

Risk profiles for different hazards may vary by country, region, species and/or production system. The application of a control and verification assurance programme based on risk should provide the necessary basis for exporting countries to certify the safety of exported food, and for importing countries to have the confidence to accept such consignments.

It is recognized that in particular developing countries may need a transition period and/or technical assistance regarding the full implementation of these Guidelines.

2. SCOPE

This guide is intended to provide the overarching principles and guidance for governments on the design and implementation of national and trade related food safety assurance programmes for residues of veterinary drugs. The current and future annexes to this guide may provide a further refinement of guidance on issues which may be relevant to the control and verification programmes for products from certain species. These annexes should be read in conjunction with the principles outlined in this guide.

3. GENERAL PRINCIPLES

Programmes for the control of residues of veterinary drugs in foods should:

i. Be based on risk using realistic risk profiles assessed as reasonably likely to be associated with food derived from the relevant productions system(s);

ii. Be prevention focussed based on the realistic risk profiles associated with the probable or known use of approved, non-approved and prohibited veterinary drugs in the production system;

iii. Include regulatory measures proportionate to the relative human health risk associated with these hazards compared with other food-associated hazards;

iv. Ensure all parties involved in the production, marketing and processing system of the animals and/or the food products derived from them are held accountable to ensure that unsafe animal products will not be sold as a result of their action or inaction;

v. Recognise that pre-harvest controls and practices are the primary means for ensuring safe food;

vi. Recognise that the primary role of audits and sampling programmes is to verify the implementation and effectiveness of the pre-harvest controls and practices;

vii. Focus on system and population based assurances; and

viii. Be cost effective and have the support of stakeholders.

It should be recognised that veterinary drugs are regulated in many countries for a variety of reasons, such as animal health, animal welfare and protection of the environment. Where these uses and the related standards do not fall under the mandate of the Codex Alimentarius Commission, they should be clearly identified and justified where, for reason of efficiency, they form part of the Competent Authority's residue control programme.

The Codex Alimentarius Commission's recommended sampling procedures for residues of veterinary drugs in food are exempted from the general sampling procedures of food commodities developed by the *Codex*

Committee on Methods of Analysis and Sampling. Accordingly, this guideline includes sampling procedures relevant for the entire control programme.

The safety of foods is achieved by the implementation of appropriate rules applied from primary production or import to retail or export and requires the participation of all parties involved. Competent Authorities should verify correct implementation of programmes and, where necessary, if action has been taken.

The reliability of laboratory results is important for the decision making of Competent Authorities. Thus official laboratories should use methods validated as fit for purpose and work under internationally accepted (e.g. ISO 17025) quality management principles.

A control programme designed and implemented according to this guideline provides reassurance for importing countries to accept consignments certified as safe by the exporting country.

4. APPROACH BASED ON RISK

An approach based on risk applied across the entire production chain and on all food groups and potential hazards will allow Competent Authorities to focus application of resources to areas of highest risk which are most likely to have an impact on consumer health protection.

Continuous application of good practices and regular control contribute more significantly to food safety than end product testing.

Residues may exert an adverse effect on consumers in a number of ways, such as:

(a) Chronic toxicological adverse effects;

(b) Acute pharmacological effects on consumers and on the microflora of the gastrointestinal tract of consumers;

(c) Allergic reactions.

Different types of controls and monitoring programme may be justified where the risk assessment identifies one or more of these other end-points as being significant for human health. Detections of non-compliant residues (e.g. those exceeding applicable MRLs) justify regulatory follow up.

Animals and/or production systems can be exposed to a variety of veterinary drugs and other chemicals that may as a result be present in the products derived from them. Their importance for consumer health protection, however, varies with type and source.

An understanding of the circumstances required for each veterinary drug input to actually pose a risk to consumers of animal products, along with an estimate of the relative likelihood of this occurring, is essential to determine the appropriate controls and verification programmes which should be included in the design of national residue control and verification programmes.

The application of a control and verification programme based on risk should provide the necessary basis for exporting countries to certify, where required, the safety of exported food, and for importing countries, subject to any additional assessment they deem necessary, to accept such consignments.

The same principles should apply to export assurance programmes as are applied to the design and implementation of national assurance programmes.

5. DEFINITIONS (FOR THE PURPOSES OF THESE GUIDELINES)

Competent Authority(ies) means the official government organisation/agency(ies) having jurisdiction[1].

Approved means officially authorised or recognised by a competent authority.

Based on risk means focussed on and proportionate to an estimate of the probability and severity of an adverse effect occurring in consumers.

Risk profiles are defined in the Codex Procedural Manual. For veterinary drugs they relate a production system to a potential consumer health risk. They are the basis for approvals and use restrictions.

System verification means obtaining overall information on the extent of application of the practices and controls.

Risk targeted verification programmes means inspection/audit and/or sampling/laboratory analysis of specific suppliers or products aimed at the detection of non-compliance.

[1] Definition used in the *Guidelines for the Production, Processing, Labelling and Marketing of Organically Produced Foods* (CAC/GL 32-1999).

Non-biased sampling refers to the random sampling of specified populations to provide information about the occurrence of residue non-compliances, typically on an annual, national basis. Compounds selected for non-biased sampling are usually based on risk profiles and the availability of laboratory methods suitable for regulatory purposes. The results of non-biased sampling are a measure of the effectiveness and appropriateness of the controls and practices within a wider segment of the production system.

Survey refers to the collection of additional data aimed at the investigation of residues linked to a specific veterinary drug use or production type.

Withdrawal time/ Withholding time (food harvest restriction) are defined in the Codex *Glossary of Terms and Definitions (Veterinary Drugs Residues in Foods)* (CAC/MISC 5-1993). A period of time may also be represented by a combination of events or other factors.

Production system means the methods or activities used to produce food for human consumption.

Quality control (in residue laboratories) means monitoring those factors associated with the analysis of a sample by a tester.

Quality assurance (in residue laboratories) means independent review to ensure that the analytical programme is performing in an acceptable manner.

Quality management system ensures that a laboratory is managed and operated in a manner that meets the requirements of an internationally recognized quality standard to produce quality data and results (e.g. ISO 17025: 2005).

6. REGULATORY FRAMEWORK

6.1 Roles

Business operators/commercial entities involved in the production, processing and marketing of food have the primary responsibility for ensuring food safety.

Competent Authorities regulate the use of veterinary drugs, verify that appropriate practices are applied and that effective measures are in place within the veterinary drug distribution and food production system to provide effective protection of consumers and facilitate trade, consistent with the goals of Codex Alimentarius.

The competent authority responsible for providing consumer assurances for foods must ensure that it has sufficient knowledge of and control over veterinary drugs that are being sold and used within the production systems and that it has sufficient knowledge of food safety.

6.2 Approval by competent authority

6.2.1 *Criteria*

Appropriate official approval criteria should be established. These criteria may include the acceptance of the assessments of other recognised competent authorities where use patterns are likely to be similar.

Approval systems should:

(a) Require an evaluation of the human safety of residues of the veterinary drug relying on a risk analysis and establishing, where appropriate, maximum residue limits;

(b) Take into account the needs of the producers in order to reduce the temptation to use unapproved veterinary drugs or prohibited substances.

Approval systems should take into account that risk profiles and management options may vary substantially among production systems and regions.

6.2.2 *Approval restrictions*

The conditions for the approval of veterinary drugs should be specified in the appropriate national regulations.

To mitigate potential risk, restrictions may be imposed on:

(a) Formulations;

(b) Criteria of use (e.g. time, species) and route of administration;

(c) Indications for use; and

(d) Withdrawal time/withholding time/food harvest restriction.

6.2.3 *National register*

All formulations of veterinary drugs approved in a country should be recorded in a national register.

6.3 Information on veterinary drugs

Information and/or education programmes on suitable use to provide effective treatment while affording protection of consumers should be provided for each approved veterinary product formulation.

6.4 Sale and use

National/regional regulations should establish which veterinary drugs may be sold domestically and how these may be used. Formulations not recorded in the national register should not be used and sanctions should be in place to act as a deterrent against such use.

It may be appropriate, where justified by a relevant risk profile for Competent Authorities, to impose additional conditions on the sale and use of certain veterinary drugs to ensure appropriate use and to prevent misuse or abuse.

Sale and use conditions may include:

(a) Requiring all sales to be subject to a prescription from a veterinarian or other professional with approved competencies;

(b) Restricting administration to individuals or professionals with approved competencies;

(c) Requiring all treated animals/production systems to be identified in specified ways;

(d) Requiring all uses to be recorded and/or notified to a unified database(s).

Efficacy and the necessity of use conditions should be regularly reviewed against the local risk profile. In doing this it should be considered that the non-availability of necessary treatments may encourage use of non-approved veterinary drugs or prohibited substances.

Competent Authorities may establish legislation/regulation that allows, as an exception, the use of non-approved veterinary drugs off-label/extra label in accordance with direct and written veterinary advice and oversight. Such legislation should be consistent with national and/or international guidance and technical information on this issue.

In animals from which milk, eggs or honey, respectively, are collected for human consumption, only veterinary drugs specifically approved for use in lactating animals, laying birds and honey bees should be used. Specific exemptions may be made for off-label/extra label use.

6.5 Responsibilities of business operators (best practice guidance)

Producers should only use veterinary drugs which have been approved for use in food producing animals. Non-approved veterinary drugs should not be used. Veterinary drugs should be used strictly in accordance with the officially approved/recognised instructions. Off-label use of veterinary drugs should only be permitted in accordance with direct and written advice from a veterinarian in accordance with national authorities' laws and regulations. Such advice should be consistent with national and/or international guidance documents and technical information on this issue.

Producers should be encouraged to seek advice of veterinarians or other competent professionals on the application of the correct withdrawal time, where the label direction for use may not be available or may not be clear.

Records should be kept of all details of the treatment and the withdrawal time/withholding time required before the animal or product from the animal can be harvested for human consumption.

Business operators (whether primary producers or others) should be required to communicate food harvesting restrictions (withdrawal/withholding times) still in place on the animal or animal product at the time of sale to subsequent purchasers of the animal(s).

Processors should be required to ensure that they only purchase and/or process animals and/or animal products from suppliers (whether primary producer or others) who can credibly attest to the suitability/safety of the animal or animal product for the purpose intended.

Producers should have appropriate on-farm food safety assurance measures in place with respect to the use of and/or exposure of food-producing animals to veterinary drugs. All workers directly involved with the animals should be familiar with these measures.

Producers should be able to identify all food-producing animals, or lots of these animals, which have been treated with or exposed to veterinary drugs to ensure compliance with withdrawal/withholding times.

Continuous food safety assurance measures such as record keeping should ensure that products (e.g. milk, eggs, honey) are harvested only if appropriate withdrawal/withholding times have been followed.

Treated or exposed animals for which the withdrawal time/withholding time has not elapsed should be kept separate from animals that have not been treated, or be positively identified to reduce the potential for mistakes.

Products from animals under harvest restrictions should be handled in such a way that ensures their product does not mix with that being harvested for human consumption. Any equipment likely to be contaminated should be adequately cleaned prior to being used on other animals.

7. VERIFICATION PROGRAMMES

7.1 Purpose

A verification programme that combines audits/inspection of various control points and point of harvest testing should be implemented. This approach will reduce reliance on chemical analyses and provide a higher degree of assurance.

The overall objective of the verification programme is to provide an appropriate degree of confidence that the practices and controls in place are adequate and being applied to the extent necessary to ensure the health of consumers of animal products. It will therefore attempt to ensure that exposure to residues in excess of the ADI rarely occurs.

Verification programmes may contribute to the:

- (a) Verification of assumptions made in the registration process;
- (b) Identification of unacceptable production, marketing and/or chains of advice;
- (c) Evaluation of the effectiveness of veterinary drug label information as it relates to food safety;
- (d) Evaluation of the effectiveness of education or risk reduction programmes;
- (e) Evaluation of Quality Management Systems;
- (f) Verification of implementation and effectiveness of corrective actions.

7.2 General design principles

Verification programmes should cover, as appropriate, the entire food chain. A combined system of inspection/audits and sampling/laboratory analysis should be implemented. The frequency, point and type of activity should be based on an assessment of the risk to provide the most effective control.

Verification programmes can be classified as follows according to objective and criteria applied to the sample selection:

- (a) System verification programmes;
- (b) Risk-targeted verification programmes;
- (c) Surveys;
- (d) Port of entry testing programmes.

Verification programmes may focus on assessing the

- (a) Effectiveness of a control system; and/or
- (b) Compliance by individuals or groups.

7.3 System and targeted verification programme design

Verification programmes should:

- (a) Define their purpose;
- (b) Identify the population being sampled;
- (c) State whether the sampling is non-biased or targeted (directed); and
    - base the number of samples for non-biased sampling protocols on statistics;
    - pre-determine targeting criteria to direct sampling;
- (d) Pre-determine the criteria to be applied to the analysis of the results;

(e) Define sampling and identification procedures that allow tracing each sample back to its origin and independent confirmation of the finding in case of dispute.

7.4 Risk Profiling

It is the responsibility of the Competent Authorities to determine the risk profiles for their country and/or production system.

The frequency and intensity of verification or inspection/audit of each drug residue chosen to be monitored under the system verification programme should depend on the veterinary drug and use profile.

Risk profile considerations concerning veterinary drugs include:

(a) The type of hazard presented;

(b) The class and severity of the adverse human health effect associated with the residue (e.g. chronic toxicity, acute pharmacological, allergic reaction, or microbiological disturbance);

(c) The use and/or production circumstances required to produce residues and the likelihood of these occurring in foods derived from the production system at concentrations and in frequencies presenting a risk to consumer health;

(d) The dietary consumption required for the residue to give rise to a realistic consumer health risk.

Competent Authorities should attempt to make realistic estimates of the types, quantities and use patterns of veterinary drugs in their jurisdiction.

Subsequently the following should be considered:

(a) Circumstances required for each veterinary drug to cause an adverse health impact on consumers;

(b) Likelihood of such circumstances occurring.

When considering and ranking the residues associated with the veterinary drugs likely to be present at some stage in the production system potential sources and exposure pathways should be described.

The following sources of veterinary drug residue should be considered:

(a) Veterinary drugs authorised in the jurisdiction of the Competent Authority;

(b) Veterinary drugs that are known to be, or suspected of being misused.

The exposure pathways of veterinary drug residue should be considered:

(a) Intended e.g. direct administration to the animals;

(b) Indirect administration to the animals through addition to feed or water;

(c) Unintended contamination via e.g. feed, water, or the environment.

Competent Authorities should, as appropriate to the risk profiles in the country and/or production system, consider the following potential pre-harvest control points for audit/inspection in the verification programme:

(a) The sellers and purchasers of veterinary drugs to verify what is being sold and how they are being marketed;

(b) The users of veterinary drugs (including farmers, veterinarians and feed compounders) to verify how drugs are actually being used in the production systems, e.g. according to label, what records are being kept and how the treatment status of animals is identified;

(c) The animal and animal product distributors to verify that any food harvest restrictions associated with the animal or product are effectively communicated;

(d) The assurance systems used by processors and/or producers to ensure the suitability of the animals or product they are being supplied with for the purposes they intend using it for.

8. CHOICE OF VERIFICATION PROGRAMME

8.1 System verification programmes

In setting up system verification programmes the following should be considered:

(a) Examination of the relevant control points of the control system;

(b) Non-biased sampling of a specified population with broadly similar attributes so that the results can be used to derive a statistical confidence as to the extent of control present in that population as a whole.

System verification programmes can focus on the degree of application of specific controls in the process or can focus on monitoring the residues in the animals/products at or close to the point of harvest.

Non-biased sampling programmes should be used in order to find out whether one of the controls within the system needs adjusting. They should not be relied upon for product evaluation.

Where the Competent Authority has linked the approval of a veterinary drug to particular use conditions/restrictions in order to avoid misuse or abuse, the appropriateness of the use conditions/use restrictions should be regularly verified with risk-targeted verification programmes as to their efficacy and necessity to manage the risk posed by the use of the veterinary drug.

Generally non-biased sampling protocols are not efficient in detecting low incidences of non-compliance. Where such incidences are a potential significant risk to human health other assurance programmes should be employed.

8.2 Risk targeted verification programmes

In setting up risk targeted verification programmes the following should be considered:

(a) Previous performance, history of non-compliance;

(b) The quality management components usually relied on;

(c) Potential risk factors which may be correlated with an increased use of veterinary drugs such as;

- high somatic cell counts in milk, or
- significant ante- or post-mortem findings e.g. injection site lesions or resolving pathology;

(d) Any other information linked to non-compliance and drug use.

Competent Authorities may complement the risk-targeted pre-harvest verification programmes with established risk-targeted post-harvest verification programmes.

8.3 Surveys

Surveys may be performed to:

(a) Assess the initial situation before a verification programme is started;

(b) Evaluate the efficiency and appropriateness of specific aspects of control programmes;

(c) Monitor the impact that variables, such as location, season, or age, may have on the presence, absence or concentration of a residue.

8.4 Review

Control and verification programmes should be regularly reviewed to ensure their continued efficacy and/or necessity, as well as to review the potential impact of changes to the risk profiles.

Where a significant incidence of non-compliance is identified in any one year and consequent changes to the control programme implemented, a higher standard of verification may be appropriate until the effectiveness of the corrective actions has been demonstrated. Some of the selected lower risk profile veterinary drugs should be considered for rotation in and out of the programme based on history of compliance to ensure that the scope is as wide as possible.

9. SAMPLE TAKING

9.1 General principles

Appropriate mechanisms to prevent possible bias occurring in both the selection and taking of samples should be put in place.

Ideally, samples should be taken before animals and/or products are commingled with animals or product from other suppliers.

9.2 Traceability/product tracing

Competent Authorities should ensure that all samples can, throughout the sampling, storing, shipping, analysis and reporting, be traced back to their origin.

Each sample needs to be clearly identified so that appropriate follow-on actions can be applied in case of non-compliant results.

If sub-units of a consignment are sampled, care should be taken to identify those sub-units clearly. Sufficient sample should be taken to allow for unprocessed sub-units to be retained allowing possible independent confirmation of the findings.

10. STATISTICAL CONSIDERATIONS

10.1 General

The number of samples for system verification programmes can be statistically pre-determined (*see* Appendix A for additional guidance).

In designing a sampling protocol it is essential to define both the purpose of the programme and the population of interest. It is also important to define the criteria to be applied when analysing the results with respect to the need/desirability for any further action, and especially how such criteria and actions directly relate to the protection of human health.

Ultimately “a population” made up of “units of food consumed” is the most relevant to human health. However, as it is the application of appropriate pre-harvest practices and controls which ensures food safety, a sampling strategy which verifies both the appropriateness and extent of compliance of these pre-harvest practices and controls can be used to provide appropriate assurances that the health of consumers is unlikely to be negatively affected. Generally the population of interest for targeting pre-harvest compliance/appropriateness verification information will be those population units to which common practices and controls should be applied such as:

(a) The seller of the veterinary drug input into the production system;

(b) The producer;

(c) The supplier of the animals or animal product to the processor; or

(d) The processor.

However, because the potential consequences to human health are much larger when large production units (farms) are out of control, the usual pre-harvest population randomly sampled is a standardised unit of production sold at any one time e.g. individual animal, vat of milk, barrel of honey, or defined weight of aquaculture product. In this way the larger producers/suppliers should effectively have a greater probability of being sampled while still maintaining the randomness of the sampling protocol.

Generally, conclusions will be drawn from the prevalence, or lack thereof, of non-complying results in the units sampled during the production season or calendar year. However, where problems are found during the course of the production season, corrective actions may have already been applied and have started to have a positive effect well before the end of production season or calendar year. For small populations, or for either low risk or reasonably stable exposure scenarios, several production seasons or calendar years may be used/ needed to collect the number of samples statistically determined to give the required confidence.

Where it is possible to further refine and describe the affected population associated with defined risk factors such as season, region or specific type of production, then a correlation of the sampling protocol to such a co-variable may be justified.

The point at which a sample is taken depends on the objective of the specific programme. Where the objective is to verify the effectiveness of controls at the supplier stage, generally samples are taken at the point of sale/harvest in order to correlate the unit sampled with a supplier or producer.

On-farm sampling may also be used as part of a pre-harvest quality assurance programme or where there are concerns associated with the possible use of substances prohibited by the Competent Authority.

Where the objective is to verify the overall effectiveness of a system at ensuring the general population’s exposure is less than the ADI then multiple sample units can be combined before analysis, or commingled product sampled and analysed.

Where the objective is to verify the credibility and effectiveness of the control and verification programmes present in an exporting country, samples may be taken from standardised units of export at the port of entry. Such secondary verification programmes have quite different design considerations with respect to their objective, the population of interest and the type of response to any identified incidence of non-compliance. The statistical tables in Appendix A are not relevant to such programmes and the number of samples should reflect the importing country’s confidence in the performance of the exporting country.

10.2 Retention of consignments during laboratory analysis

Competent authorities should not routinely retain lots of production associated with randomly selected samples pending the availability of the analytical results. Competent Authorities may routinely retain lots of production where it is considered likely that a risk targeted test will produce non-compliant results that present a potential risk for consumer health.

10.3 Result interpretation

A greater degree of assurance is achieved if verification programmes such as statistically based systems involving non-biased sampling and risk targeted (e.g. specific suppliers or products) are operated in parallel.

The results of risk targeted verification programmes alone do not allow conclusions on the exposure of the general population with residues of veterinary drugs.

Conclusions on the exposure of the general population can be drawn from the combining the results of:

(a) Statistically based system verification programmes involving non-biased sampling; and

(b) Risk targeted verification programmes.

10.4 Port of entry testing programmes (specific requirements)

Competent Authorities should consider port of entry testing programmes only as a secondary system verification tool.

The matrices used in port of entry programmes may vary from those used for national verification programmes.

Except where a risk to health is suspected or detected, certified product should be subjected to non-biased sampling and release programmes at a frequency determined by the importing country based on the exporting country's record of compliance. Consignments of animal products tend to be heterogeneous by nature and will often be made up from a variety of animals, farms and processing dates. Results will reflect the performance of the national control and verification system as a whole and should not be extrapolated to specific judgements on other units within the consignment except where a common pre-harvest risk factor is shared and a direct health threat is indicated.

The application of directed or targeted sampling in port of entry sampling programmes is only appropriate where it is known or suspected that products share the same risk profile.

However, following the detection of non-compliant results during port of entry programmes, importing countries may increase the overall frequency of testing of directly related food of animal origin from the exporting country for a period as an added verification of the effectiveness of any additional controls being implemented by the exporting country.

In the interpretation of laboratory results of consignments of animal products it should be considered that these are made up of commingled product from a variety of animals, farms and processing dates and, therefore, heterogeneous. Because of this, results should not be taken to judge other units of a consignment except where units share a common pre-harvest risk factor and where a direct risk to health is suspected or detected.

Results of port of entry testing programmes should only be communicated if confirmed with methods fully validated for the specific matrix and analyte.

Laboratory reports on non-compliant results should include:

(a) A description of the method used;

(b) Performance characteristics of the method of analysis (including the confidence interval of the result).

Laboratory reports on non-compliant results should be distributed to all parties affected by the result (e.g. the owner of the consignment and the certifying competent authority of the exporting country).

Competent Authorities of importing countries should provide exporting countries regularly with the results of their verification programmes including information for purposes of traceability/product tracing.

In cases of non-compliance with the food safety parameters, competent authorities from the exporting country should conduct a trace back, apply appropriate corrective actions and then provide a summary of these to the importing country.

Where the type, incidence and/or frequency of non-compliance detected raises concerns as to whether the imports are meeting the standard of human health protection required by the importing country, then additional assurances may be requested.

The importing country may also choose to increase the frequency of port of entry verification to confirm that the assurances given are in fact addressing the problem.

Where residues of substances that should not be used in food producing animals in either the exporting or the importing country are detected in port of entry testing, both competent authorities should co-operate in order to identify potentially similarly affected food of animal origin and to resolve any potential wider control problem.

Resolution of such problems will require the originating country to conduct an analysis to determine the possible source of such residues, the identification of deficiencies within the country's own control and monitoring system, and subsequent application of appropriate additional controls and measures to address the situation.

In cases where the exporting country is a less developed nation, consideration should be given by the importing country to the provision of technical assistance to help resolve the issue.

The application of new sampling and testing methods may reveal the presence of types and concentrations of residues previously unknown to exist by one or both parties. The determination of the source of such residues and their significance may take some time.

Where the presence of such residues is associated with previously accepted production practices, the implementation of changes, should these be deemed necessary, may require an extended period of time for capacity building.

11. REGULATORY ACTION

11.1 Investigation of non-compliances

Competent authorities should investigate each non-compliant result to ascertain the contributing factors which lead to its occurrence and the systemic significance of the identified case.

An attempt should be made to identify the substances and the consumer health significance of their occurrence in food.

When an animal tissue/food contains residues in excess of the relevant MRL at the point of harvest the following possibilities should be considered:

(a) The veterinary drug was not used according to label or prescription instructions;

(b) A non-authorised veterinary drug or formulation was used;

(c) The recommended withholding time was not observed or is not appropriate;

(d) Treated and non-treated animals were commingled;

(e) Unintended exposure to feed, water or contaminated environment occurred;

(f) The food is part of the statistically predictable small percentage of animals with residues in excess of the MRL even when the required withdrawal period has elapsed;

(g) Sample contamination, analytical method problems or analytical error.

Laboratories should report all suspect positive samples which they have not been able to positively confirm using established confirmation criteria. This will allow competent authority to identify possible patterns of non-compliance.

11.2 Measures in case of non-compliance: Conduct

Competent authorities should adjust the scale and type of response to identified non-compliances to the relative importance that the respective hazard has for consumer health protection.

Competent authorities should take proportionate action when considering whether the non-compliance is the result of negligence or intent.

Competent authorities should in case of isolated mistakes due to ignorance or negligence require that appropriate advice and training measures are followed.

In the case of proven negligence or intent punitive measures in line with the Codex member's penal system should be considered (e.g. condemnations, fines, movement controls, etc.) to act as a deterrent.

Competent authorities should, in case of widespread non-compliance, advise stakeholders and motivate the respective business sector to initiate the necessary changes.

Competent authorities should verify that appropriate corrective action is taken and monitor the success of these measures through inspection/audits and/or sampling/laboratory analysis.

11.3 Measures in case of non-compliance: Product

Unsafe product should not be passed as fit for human consumption and should be disposed of appropriately.

Where the results of samples taken on farm for risk targeted verification programmes do not provide the necessary confidence that the rest of the lot has been produced using appropriate practices and controls, the lot should not be passed for human consumption until sufficient information can be generated to provide the required degree of assurance as to its safety.

Where the results indicate there is a direct risk to consumer health, an attempt should be made to trace and remove all similarly affected products.

In non-biased sampling programmes the unidentified proportion may represent a much greater potential threat to consumers than the identified proportion. Accordingly, any actions taken with respect to the identified non-compliant lot are less significant than the actions taken on the system as a whole.

When pre-harvest controls are not carried out or are unreliable due to a high incidence of misuse of veterinary drugs, more frequent post-harvest verification may be appropriate to provide the required degree of consumer assurance. This should be regarded as an interim measure only until the appropriate corrective actions to the control programme have been put in place and subsequently demonstrated to be effective.

11.4 Corrective action in case of non-compliance

Depending on the results of such investigations local and/or systemic corrective actions may be considered appropriate to prevent reoccurrence.

Where the investigation of non-compliances indicates that use and distribution provisions for the substance(s) are inappropriate, competent authorities should take appropriate corrective action by modifying approval and distribution rules.

Where the investigation of non-compliances identifies local or systemic control failures, competent authorities should ensure that appropriate corrective action is taken at the relevant points.

The competent authority should verify that the measures are taken. Respective action should be proportionate in time and intensity to the consumer health hazard, scale and frequency of the non-compliance.

In cases where the failure lies outside of the direct control of the business operator the competent authority should prevent repetition of the failure by applying appropriate measures at the relevant control point.

12. INTERACTION BETWEEN THE CONTROL PROGRAMMES OF TWO COMPETENT AUTHORITIES

Competent authorities should co-operate to ensure consumer health in all countries is protected.

This co-operation aims at achieving better assurance than can be achieved through sole reliance on port of entry inspection programmes.

Trading countries should exchange copies of their control and verification programmes along with the results of these programmes from preceding years on a regular basis.

In order to facilitate trade from developing countries longer transition periods and technical assistance regarding all aspects of a residue control programme should be considered.

ANALYTICAL METHODS FOR RESIDUE CONTROL

General Consideration on Analytical Methods for Residue Control

13. INTRODUCTION

Analytical methods used to determine compliance with maximum residue limit for veterinary drugs (MRLVDs) should be suitable for routine use by competent authorities of member governments for their testing programmes for all residues of veterinary drugs and substances which may be used as veterinary drugs. This includes certain pesticides which have veterinary uses and that may be present as residues in commodities. These methods may be used for the analysis of randomly selected survey samples in a national regulatory control programme to determine compliance with established MRLVDs, for the analysis of

targeted samples when there is reason to suspect non-compliance with MRLVDs or for the collection of data for use in estimation of intake.

Methods may also be required in regulatory control programmes for the detection of residues of substances for which ADIs and MRLVDs have not been established by the Codex Alimentarius Commission. For some substances, the toxicological evaluation leads to the conclusion that an ADI or MRLVD should not be established. For such substances, the determination of the lowest concentration at which the residue can be detected and the identity confirmed in a food is a primary concern in the method validation. Performance characteristics related to quantitative analyses may be less critical for such substances, where detection and confirmation of the presence of the substance as a residue is the major issue. Confirmation of identity of a residue is generally based on the comparison of a set of characteristics of a detected substance with those of a known standard of the suspected residue.

Suitably validated methods are not always available for all possible combinations of veterinary drug residues and foods. Competent authorities responsible for designing national residue control programmes should ensure that appropriate residue methods of analysis are used to assure compliance with Codex MRLVDs. This may sometimes require the development and validation of a new analytical method or the extension of the validation of an existing analytical method to include a new combination of analyte and matrix. Appropriate regulatory action may then be taken against adulterated products, consistent with the reliability of the analytical data.

14. INTEGRATING ANALYTICAL METHODS FOR RESIDUE CONTROL

Analytical methods for veterinary drug residues in foods must reliably detect the presence of an analyte of interest, determine its concentration and correctly identify the analyte. When residues resulting from the use of approved veterinary drugs are detected at concentrations above an established MRLVD, the results should be confirmed before regulatory enforcement actions are taken. In the case of substances which have been banned from use in food-producing animals by a competent authority, or for which an ADI and MRLVDs have not been established for toxicological reasons, the confirmed presence of residues at any concentration in a food may result in regulatory action.

The principal performance attributes of analytical methods used in residue control programmes are dependent on whether a method is intended to simply detect, to quantify, or to confirm the presence of a target residue. Completion of a full collaborative study[2] is not a requirement for recognition of a method to be placed in one of these three categories.

Screening methods are qualitative or semi-quantitative in nature and are used as screening methods to identify the presence (or absence) of samples from a herd or lot which may contain residues which exceed an MRLVD or other regulatory action limit established by a competent authority. These methods may not provide adequate information to accurately define the concentration present or, to confirm the structure of a residue but may be used to quickly determine which products require further testing and which can be released. They may be applied to a sample at the point of entry into the food chain, site of inspection or on receipt of a sample at the laboratory to determine if the sample contains residues which may exceed a regulatory limit. Such methods usually provide greater analytical efficiency, can sometimes be performed in non-laboratory environments and may be less expensive for use in regulatory control programmes than tests conducted within a laboratory. Use of screening methods allows the laboratory resources to be focused on analysis of the presumptive positive (suspect) samples identified using this test. These methods, which should have a defined and low false negative rate, should not be used alone for residue control purposes on official samples without the availability of suitably validated quantitative and/or confirmatory methods to apply to any samples identified as potentially not in compliance with an MRLVD.

Quantitative methods provide quantitative information which may be used to determine if residues in a particular sample exceed an MRLVD or other regulatory action limit, but do not provide unequivocal confirmation of the identity of the residue. Such methods which provide quantitative results must perform in good statistical control within the analytical range that brackets the MRLVD or regulatory action limit.

Confirmatory methods provide unequivocal confirmation of the identity of the residue and may also confirm the quantity present. Confirmatory methods are the most definitive and frequently are based on combined chromatographic and mass spectrometric techniques, such as liquid chromatography – mass spectrometry (LC/MS). Such methods when used for confirmation of residue identity should provide reliable structural information within established statistical limits. When the confirmatory method does not provide quantitative information, the quantification result of the original quantitative method should be verified by analysis of

[2] Horwitz, W. 1995. Protocol for the design, conduct and interpretation of method performance studies. *Pure and Applied Chemistry*, **67:**331-343.

replicate test portions using the original quantitative method or a suitably validated alternative quantitative method.

These three categories of methods – screening, quantitative and confirmatory - often share some performance characteristics. In addition, each category has other specific considerations. Understanding the relationship between these three categories of methods is important in the development and operation of a balanced residue control programme. These three categories of methods may be applied sequentially in a residue control programme.

Samples which test "positive" with the screening method are considered as suspect and are usually designated for further laboratory testing using more definitive methods. This could include repeat testing of replicate test portions with a screening method, but typically quantitative and/or confirmatory methods are used in the laboratory to establish that the sample does contain residues in excess of the regulatory limit. Such tests should be conducted on new test portions of the sample material used in the initial screening test to confirm that the analyte detected in the initial test is definitely the suspected compound and that the MRLVD (or other regulatory action limit established by the authority) has indeed been exceeded. The performance attributes, or characteristics, which must be determined during method validation for each type of method – screening, quantitative, confirmatory – are presented in the Chapter "*Attributes of Analytical Methods for Residues of Veterinary Drugs in Foods*" below.

15. CONSIDERATION FOR SELECTION AND VALIDATION OF ANALYTICAL METHODS

15.1 Identification of Methods Requirements

15.1.1 *Method scope*

The intended purpose of the method is usually defined in a statement of *scope* which defines the analytes (residues), the matrices (tissues, milk, honey, etc.) and the concentration range to which the method applies. It also states whether the method is intended for screening, quantitative, or confirmatory use. The Competent Authority must establish an appropriate *marker residue* for each drug for which an MRLVD has been established and should also designate a preferred *target tissue* to be sampled for testing.

15.1.2 *Marker residue*

The MRLVD is expressed in terms of the marker residue, which may be the parent drug, a major metabolite, a sum of parent drug and/or metabolites or a reaction product formed from the drug residues during analysis. In some cases, the parent drug or the metabolite may be present in the form of a bound residue which requires chemical or enzymatic treatment or incubation to be released for analysis. It is important that the marker residue should, whenever possible, provide unequivocal evidence of exposure to the drug. In rare situations, it is necessary to use compounds as marker residues which may also result from sources other than exposure to the drug. In such cases, additional information is required to ascertain the probable source of the residue is exposure to the drug. An example of such a situation is the use of semi-carbazide, which may occur from other sources, as a marker residue for the drug nitrofurazone.

15.1.3 *Target Tissue*

The usual target tissue selected by competent authorities to be tested for veterinary drug residues in a residue control programme is the edible tissue in which residues of the marker residue occur at the highest concentrations and are most persistent. For lipophilic substances, the usual target tissue is fat. For most other substances, the target tissue is liver or kidney, depending on the primary route of elimination. One of these tissues is usually the target tissue designated for use in testing of domestically produced foods of animal origin. The organ tissues may not be available for testing imported products, so muscle tissue may be the target tissue for testing of these commodities. In some cases, such as drugs which are normally administered as injectable formulations, testing of muscle tissue from suspected injection sites may be required. The regulatory programme manager and the laboratory managers need to clearly identify the testing objectives and the analytical requirements required in terms of target tissues, marker residues and concentration ranges to ensure suitable methods are used in the regulatory control programme. In certain situations, Competent Authorities may also use biological fluids such as urine or serum to indicate the presence or absence of residues of interest.

15.2 Implementing other Codex Alimentarius Commission Guidelines

The Codex Alimentarius Commission has issued a guideline for laboratories involved in the import/export testing of foods[3] which recommends that such laboratories should:

(a) Use internal quality control procedures, such as those described in the "Harmonised Guidelines for Internal Quality Control in Analytical Chemistry laboratories[4]";

(b) Participate in appropriate proficiency testing schemes for food analysis which confirm to the requirement laid out in "the International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories[5];

(c) Comply with the general criteria for testing laboratories laid down in ISO/IEC Guide 17025:2005 "General requirements for the competence of calibration and testing laboratories"; and

(d) Whenever available, use methods which have been validated according to the principles laid down by the Codex Alimentarius Commission.

Methods used for analyses of veterinary drug residues in foods should be capable of detecting the compounds included in the residue control programme. The analytical recovery and precision for the target foodstuffs should meet the criteria stated elsewhere in this document. The methods should be used within an established laboratory Quality Management System which is consistent with the principles in the document on internal quality control referenced above. When methods which have not been subjected to a multi-laboratory performance trial are used in a regulatory programme for control of veterinary drug residues in foods, the quality control and quality assurance procedures applied with these methods require careful definition, implementation, and monitoring. In the case of methods which have been through multi-laboratory trials, performance characteristics, such as recovery and precision, are defined through the results obtained during the study. For a method validated within a single laboratory, data must be generated to define the performance characteristics expected of the method when used by analysts within that laboratory. The on-going performance must be monitored through the Quality Management System in place in the laboratory.

15.3 Method Validation and Fitness for Purpose

The process of method validation is intended to demonstrate that a method is *fit-for-purpose*. This means that in the hands of a properly trained analyst using the specified equipment and materials, and following the procedures described in the method, reliable and consistent results can be obtained within specified statistical limits for the analysis of a sample. The validation should address the issues of marker residue, target tissue and concentration range identified by the laboratory in consultation with the residue programme manager. When the method protocol is followed, using suitable analytical standards, results within the established performance limits should be obtained on the same or equivalent sample material by a trained analyst in any experienced residue control laboratory.

Multi-laboratory method performance studies generally satisfy the analytical requirements for use in a regulatory programme. These methods are subjected to a properly designed inter-laboratory study with analysts in independent laboratories, so that different sources of reagents, materials, and equipment are used by the participants.

Quantitative methods studied collaboratively according to the revised harmonized protocol adopted in 1995 by AOAC International, the International Union of Pure and Applied Chemistry (IUPAC), and the International Standards Organization (ISO) have been evaluated in a minimum of 8 laboratories, unless highly complex equipment or other unusual requirements were identified (in such cases, a minimum of 5 participating laboratories is required)[5]. Collaborative studies of qualitative methods currently require a minimum of 10 participating laboratories. Collaborative studies conducted prior to 1995 completed method evaluation in a minimum of six laboratories in an acceptable, statistically designed study. These multi-laboratory method performance studies generally satisfy the analytical requirements for use in a regulatory programme, as information on method performance in the hands of different analysts in different laboratories is obtained through these studies. However, relatively few of the analytical methods currently used in residue control programmes for veterinary drug residues in foods have been validated by such a multi-laboratory study. Collaborative study designs are based on the analyses of coded duplicate test materials which represent the combinations of analytes, matrices, and concentrations included in the scope of the method and include an independent peer-review of both the study design and the results. In some situations, multi-laboratory studies may be conducted which do not have the minimum number of laboratories required to qualify as a collaborative study. Such studies, when conducted using the same scientific principles of design, evaluation,

---

[3] CAC/GL 27-1997. *Guidelines for the Assessment of the Competence of Testing Laboratories Involved in the Import and Export Control of Food.*

[4]. Pure *and Applied Chemistry,* **67 (1995)**: 649-666.

[5] *Pure* and *Applied Chemistry,* **78** (2006) 145-196.

and review as are applied in collaborative studies, can provide useful information on method performance in the hands of multiple analysts in different laboratories, but do not provide the same degree of statistical confidence obtained from the results of a collaborative study.

Multi-laboratory and collaborative studies of methods usually do not encompass all possible combinations of residue, tissue and species to which the method may subsequently be applied. Methods may be extended to include related analytes, additional tissues, species or products (or combinations of these not included in the original multi-laboratory study) by completing additional within-laboratory studies. Analytical results from method extension studies may require additional review before use in a regulatory programme. Whenever possible, analytical results obtained using methods that have not been validated by traditional inter-laboratory study should be compared with results obtained using a method which has been validated through a collaborative or multi-laboratory study or tested using sample materials from a recognized proficiency programme. The comparison should be based on a statistically acceptable study design using portions of the same (homogeneous) samples. The data from such studies should be independently reviewed by a qualified third party (such as a QA unit, a peer group of regulatory scientists, auditors of national accreditation body) to determine the comparability of method performance.

Some residue control methods that have been demonstrated to be suitable to determine compliance with MRLVDs have a history of use in one or more expert laboratories, but have not been subjected to a formal multi-laboratory study. These methods were demonstrated to be suitable at the time of initial regulatory use and have continued in use over an extended period of time either in the absence of alternative validated methods, or because they remain a preferred choice for reasons which may include use of available technology, cost, reliability and suitability for use within the constraints of a national programme. Although evidence of a formal collaborative or multi-laboratory method trial is lacking, the method performance has been demonstrated through successful use and from quality control data in one or more laboratories over time.

Most regulatory laboratories rely on the use of veterinary drug residue methods which have not have been subjected to a multi-laboratory study. Factors which have contributed to this situation include a requirement for specialized expertise or equipment, cost of such studies, lack of suitable collaborating laboratories, analyte and/or sample instability and rapidly changing technologies. While for many years the focus on equivalency of analytical results was based on the use of standardized methods which had performance characteristics defined based on collaborative study, accredited laboratories now operate in an environment where it is the responsibility of the individual laboratory to demonstrate that the methods used and the analytical results produced meet performance criteria established in consultation with a client. In the absence of methods validated through inter-laboratory method trials, regulatory laboratories must frequently use analytical methods which have been subjected to studies conducted within their own laboratory to characterize the method performance.

15.4 Single Laboratory Validation – The Criteria Approach

A guidance document on single laboratory validation of methods, "Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis", has been published as a technical report by the IUPAC[6]. The Procedural Manual[7] recognizes that inter-laboratory validated methods are not always available or applicable, particularly for multi-analyte/ multi-substrate methods and new analytes. In such cases, methods may be validated in a single laboratory to meet the General Criteria for the Selection of Methods of Analysis, as well as the additional criteria:

(a) The method is validated according to an internationally recognized protocol (for example, the IUPAC Guidelines for Single Laboratory Validation of Methods of Analysis, referenced above);

(b) Use of the method is embedded in a Quality Management System in compliance with the ISO/IEC 17025 (2005) Standard or with the Principles of Good Laboratory Practice;

(c) The method should be complemented with information on accuracy demonstrated for instance by:

- regular participation in proficiency schemes, where available;
- calibration using certified reference materials, where applicable;
- recovery studies performed at the expected concentration of the analytes;
- verification of result with other validated method where available.

---

[6] Thompson, M., Ellison, S.L.R. & Wood, R. (2002) Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis. *Pure and Applied Chem*istry **74**: 835-855.

[7] FAO/WHO Codex Alimentarius Commission Procedural Manual.

The criteria approach, which combines a single laboratory validation model with a requirement that methods meet specific performance specifications, has been adopted by some regulatory authorities.

Attributes of Analytical Methods for Residues of Veterinary Drugs in Foods

16. INTRODUCTION

The performance characteristics of analytical methods used to determine compliance with MRLVDs must be defined and proposed methods evaluated accordingly. This will assure reliable analytical results and provide a secure basis for determining residues of veterinary drugs in foods for commodities in international trade. The chapter "*General Considerations of Analytical Methods for Residue Control*" above, presents a discussion of general types or categories of regulatory methods, and provides a scheme for using these analytical methods based upon their intended purpose in a regulatory framework. In the discussion below, attributes common to the three categories of methods (referred to as Confirmatory, Quantitative and Screening methods) for determining compliance with Codex MRLVDs are presented. The additional attributes that are applicable to only one or two categories of methods are also discussed.

17. METHOD DEVELOPMENT CONSIDERATIONS

The development of an analytical method requires analysts experienced in the analytical techniques to be used, as well as appropriate laboratory space, equipment, and financial support. Before initiating method development activities, the intended use and need for a method in a residue control programme should be established, including the required performance parameters. Other considerations include the required scope of the method (compound or class of compounds of interest and types of sample materials), potential interfering substances, the required performance characteristic of the measurements system, the pertinent physical and chemical properties that may influence method performance, the specificity of the desired testing system and how it will be determined, analyte and reagent stability data and purity of reagents, the acceptable operating conditions for meeting method performance factors, sample preparation guidelines, environmental factors that may influence method performance, safety considerations, and any other specific information pertinent to programme needs. In particular, stability of standards, both under normal conditions of storage and use and during processing of samples, should be assessed. Analyte stability in samples during typical conditions of sample storage prior to analysis should also be determined, including any period for which a sample may be held pending a potential re-analysis for confirmatory purposes.

Establishing method performance attributes is essential, as these provide the necessary information for food safety agencies to develop and manage their public health programmes. Performance attributes for analytical methods also provide a basis for good management decisions in future planning, evaluation, and product disposition. For the animal health care industry, it provides a guideline for knowing exactly what performance must be achieved in developing analytical procedures. All will benefit by having well defined analytical method performance factors. Method performance requirements will vary, depending on whether the method is used for the screening, quantification, or confirmation of a residue for which Maximum Residue Limits have been established, or for residues of a drug for which an ADI and MRLVDs have not been recommended. In the latter case, the Competent Authority may establish a minimum performance standard which must be met by analytical methods used for regulatory control purposes. However, when no safe concentrations of these compounds in foods have been established, the Competent Authority may review such limits periodically to ensure they reflect improvements in technology and analytical capability. When such limits have not been formally established by the Competent Authority, they are usually established *de facto* by the detection capabilities of the methods used in the regulatory laboratories.

18. ANALYTICAL PERFORMANCE CHARACTERISTICS

18.1 Performance Characteristics of Screening Methods

Screening methods are usually either qualitative or semi-quantitative in nature, with the objective being to discriminate samples which contain no detectable residues above a threshold value ("negatives") from those which may contain residues above that value ("positives"). The validation strategy therefore focuses on establishing a threshold concentration above which results are "positive", determining a statistically based rate for both "false positive" and "false negative" results, testing for interferences and establishing appropriate conditions of use.

For a screening test, particularly those involving test kit technologies, the term "*sensitivity*" refers to the lowest concentration at which the target analyte may be reliably detected within defined statistical limits. In the AOAC Performance Tested Program ™ for test kits, this is determined experimentally by testing a minimum of 30 residue-free sample materials fortified with the analyte at the target concentration. The sample materials should be from at least six different sources (that is, at least 5 replicates from each of at

least 6 sources), all of which should yield a positive result when fortified at the target concentration. Three or more negative results constitute a failure of the sensitivity test. If one or two of the results are negative, the experiment should be repeated and two negative results would then constitute failure. The experiment should be repeated with known incurred material at the target concentration, if such material is available.

The "*selectivity"* of a screening method refers to the ability of the test to determine that samples which give a negative response are truly negative. The test must also be able to distinguish the presence of the target compound, or group of compounds, from other substances which may be present in the sample material. It normally is not as great as that of a quantitative method, because screening methods often take advantage of a structural feature common to a group or class of compounds. These methods, which generally fit into the screening methods category, are often based on microbiological growth inhibition, immunoassays, or chromogenic responses which may not unambiguously identify a compound. The selectivity of a screening method may be increased when it is used as a detection system after chromatographic or other separation technique. To demonstrate a selectivity rate of at least 90% with 95% confidence (which is recommended for screening tests), 30 replicate analyses are conducted on representative blank sample matrix materials from a minimum of six different sources. All results should be negative. Additional tests for potential interferences and cross-reactivity may then be conducted by testing blank matrix material fortified with potential interfering substances, such as other drugs which might be used in animal treatment, potential environmental contaminants, drug metabolites, or chemically related compounds. Again, responses should be negative when these compounds are present at concentrations which might reasonably be expected to be present in a sample.

The "cut-off" or threshold for the test for a particular compound is established by conducting concentration-response experiments, typically using 30 replicates (from at least six sources) fortified at each of a series of increasing concentrations. Once the concentrations have been established where all 30 replicates give a negative response and all 30 replicates give a positive response, the experiment is repeated using the blank matrix materials fortified at four evenly spaced concentrations between the "all negative" and "all positive" concentrations. An additional set is tested at a concentration 20% above the "all positive" concentration. Statistical analysis of the results enables the user to establish a reliable detection concentration at the required confidence level (usually 95%)[8].

18.2 Performance Characteristics for Quantitative Methods

Selectivity, the ability of an analytical method to detect and discriminate the signal response from a compound in the presence of other compounds which may be present in the sample material, is of particular importance in defining the performance characteristics of methods used in regulatory control programmes for veterinary drug residues in foods. There are two aspects which must be considered – the ability of the method to provide a signal response which is free from interferences from other compounds which may be present in a sample or sample extract and the ability of the method to unequivocally identify a signal response as being exclusively related to a specific compound. For a quantitative method, the requirement is that the signal used for quantification should relate only to the target analyte and not contain contributions for co-extracted materials. Chromatographic analyses based on peaks which are not fully resolved provide less reliable quantitative results. Use of element-specific detectors or detection wavelengths or mass-selective detectors which are more specific to a particular compound or structure, combined with chromatographic separation, improves the selectivity of quantitative methods for veterinary drug residues in foods.

In addition to the selectivity of a method, the ability of the method to provide a quantitative result which is reliable must be demonstrated. This consists of two factors:

(a) The closeness of the result to the true or accepted value for the concentration of analyte present in the sample material, expressed in terms of *accuracy*, *trueness,* or *bias*; and

(b) The ability of the method to provide consistent results on replicate determinations, expressed in terms of *precision* (*repeatability* and *reproducibility).*

It is recommended that methods used to support Codex MRLVDs should meet the performance standards for trueness and precision listed in Table 1, where CV_A refers to the coefficient of variation determined by test portions of blank matrix fortified prior to extraction and CV_L is the overall laboratory variability which includes a 10% estimate for variability of sample processing[9].

---

[8] Finney, D.J. (1978) *Statistical Method in Biological Assay, 3rd edition*. MacMillan Publishing Co., New York.

[9] Fajgelj A., Ambrus A., eds. (2000) Principles of Method Validation, Royal Society of Chemistry, Cambridge UK.

Table 1. Performance criteria which should be met by methods suitable for use as quantitative analytical methods to support MRLVDs for residues of veterinary drugs in foods[10]

Concentration µg/kg	Coefficient of Variability (CV)				Trueness
	Repeatability (Within-Laboratory, CV_A) %	Repeatability (Within-Laboratory, CV_L) %	Reproducibility (Between-Laboratory, CV_A) %	Reproducibility (Between-Laboratory, CV_L) %	Range of Mean % Recovery
≤ 1	35	36	53	54	50-120
1 to 10	30	32	45	46	60-120
10 to 100	20	22	32	34	70-120
100 to 1000	15	18	23	25	70-110
≥1000	10	14	16	19	70-110

The *accuracy* of a method may be determined by analysis of a certified reference material, by comparison of results with those obtained using another method for which the performance parameters have previously been rigorously established (typically, a collaboratively studied method) or, in the absence of reference materials or methods validated by inter-laboratory trial, by determination of the *recovery* of analyte fortified into known blank sample material. The determination of accuracy as recovery is frequently used in validation of methods for veterinary drug residues in foods, as both certified reference materials and methods validated by inter-laboratory trial are often not available. The accuracy of a measurement is closely related to *systematic error* (analytical method bias) and analyte recovery (measured as percent recovery). The accuracy requirements of methods will vary depending upon the planned regulatory use of the results. The accuracy should be carefully characterized at concentrations near the MRLVD or target concentration for regulatory action (typically at concentrations from 0.5 to 2.0 times the target concentration) to ensure that regulatory action is only taken on samples containing residues which can be demonstrated to exceed the regulatory action limit with a defined statistical confidence.

Recovery is usually expressed as the percentage of analyte experimentally determined after fortification of sample material at a known concentration and should be assessed over concentrations which cover the analytical range of the method. In interpreting recoveries, it is necessary to recognize that analyte added to a sample may not behave in the same manner as the same biologically incurred analyte (veterinary drug residue). In many situations, the amount of an incurred residue that is extracted (the yield or recovered fraction) is less than the total incurred residues present. This may be due to losses during extraction, intra-cellular binding of residues, the presence of conjugates, or other factors that are not fully represented by recovery experiments conducted with analyte-fortified blank tissues. At relatively high concentrations, analytical recoveries are expected to approach one hundred percent. At lower concentrations, particularly with methods involving extensive extraction, isolation, and concentration steps, recoveries may be lower. Regardless of what average recoveries are observed, recovery with low variability is desirable so that a reliable correction for recovery can be made to the final result, when required. Recovery corrections should be made consistent with the guidance provided by the Codex Alimentarius Commission[10].

Precision, which quantifies the variation between replicated measurements on test portions from the same sample material, is also an important consideration in determining when a residue in a sample should be considered to exceed an MRLVD or other regulatory action limit. Precision of a method is usually expressed in terms of the within-laboratory variation (*repeatability*) and the between-laboratory variability (*reproducibility*) when the method has been subjected to a multi-laboratory trial. For a single laboratory method validation, precision should be determined from experiments conducted on different days, using a minimum of six different tissue pools, different reagent batches, preferably different equipment, etc. and preferably by different analysts. Precision of a method is usually expressed as the standard deviation. Another useful term is relative standard deviation, or coefficient of variation (the standard deviation, divided by the absolute value of the arithmetic mean). It may be reported as a percentage by multiplying by one hundred.

[10] CAC/GL 37-2001 *Harmonized IUPAC Guidelines for the use of Recovery Information in Analytical Measurement*; see also Thompson, M., Ellison, S., Fajgelj, A., Willetts, P., & Wood, R. (1999) Harmonised Guidelines for the Use of Recovery Information in Analytical Measurement, *Pure Applied Chemistry*, **71:** 337-348.

Method variability, achieved in a laboratory developing a method, is usually less than the variability achieved by another laboratory that may later use the method. If a method cannot achieve a suitable standard of performance in the laboratory where it was developed, it cannot be expected to do any better in other laboratories.

Quantitative methods are usually based on a comparison of the response from an analyte in a sample with the response from standards of the analyte in solution at known concentrations. In method development and validation, the calibration curve should first be determined to assess the detector response to standards over a range of concentrations. These concentrations (a minimum of five, plus blank) should cover the full range of analytical interest and the resultant curve should be statistically expressed. However, although it is recommended practice to include a suitable blank with the calibration samples, this does not imply that it is acceptable to extrapolate into the region of the curve below the low standard to obtain a quantitative result. The analytical function relates the response for the analyte recovered from sample material at various concentrations throughout the range of analytical interest. For analytes for which an MRLVD or regulatory action limit has been established in a particular sample material (matrix), response is typically determined for known blank sample material and for blank sample material fortified at a range of concentration above and below the MRLVD (use of 6 different sources of blank materials is recommended).

The analytical function experiment data can also be used to calculate the analytical recovery at each concentration and is of particular importance when the presence of matrix co-extractives modifies the response of the analyte as compared to analytical standards. The *linearity* is determined from the analytical function experiments and is the statistical expression of the curve obtained for the analysis of sample materials fortified at the target concentrations. It is typically determined from a linear regression analysis of the data, assuming there is a linear response. It is increasingly common in methods for veterinary drug residues in foods to base the quantitative determination on a standard curve prepared by addition of standard to known blank representative matrix material at a range of appropriate concentrations which bracket the target value (the analytical function). Use of such a "tissue standard curve" for calibration incorporates a recovery correction into the analytical results obtained.

It is also necessary to establish the lower limits at which reliable detection, quantification, or confirmation of the presence of an analyte may be performed using a particular analytical method. The *detection limit* may be described in practical terms as the lowest concentration where the analyte can be identified in a sample. It can be estimated using the standard deviation ($s_{y/x}$) from the linear regression analysis of the standard curve generated in the analytical function experiment described above[11]. Using this approach, the limit of detection is calculated using the y-intercept (assuming a positive value) of the curve plus three times $s_{y/x}$. This approach provides a conservative estimate of the detection limit. The detection limit can also be estimated by measurements on representative test materials as the weakest relevant response of the analyte in the blank plus three times its standard deviation. It is often necessary to fortify test materials at a concentration resulting in a barely detectable response to obtain an approximation of the standard deviation of the blank when using this approach.

The *limit of quantification* (LOQ), also referred to as limit of quantification or quantification limit may be established from the same experiments using the y-intercept of the curve plus ten times $s_{y/x}$. For methods used to support MRLVDs established by the Codex Alimentarius Commission, the limit of quantification should meet the criteria for precision and accuracy (recovery) in Table 1 and should be equal to or less than one-half the MRLVD. However, when the limit of quantification of a method is lower than the actual concentrations monitored for compliance with a MRLVD, the validation and subsequent application of the method should be based on a *lowest calibrated level* (LCL), which is typically 0.5x the MRLVD. For use in a regulatory programme, the limits of detection and quantification are important parameters when the method will be applied to estimate exposures to residues, where there may be an interest in monitoring residues at concentrations below the MRLVD, or when conducting residue analyses for substances which do not have ADIs or MRLVDs. For monitoring compliance with an MRLVD, it is important that a LCL be included in the analysis which adequately demonstrates that the MRL concentration may be reliably determined. The LCL of a method used to support an MRLVD should not be less than the LOQ. The Codex Procedural Manual recommends the term *determination limit* under "Terms to be Used in the Criteria Approach"[7].

18.3 Performance Characteristics for Confirmatory Methods

Selectivity, the ability of the method to unequivocally identify a signal response as being exclusively related to a specific compound, is the primary consideration for confirmatory methods**.** Certain instrumental techniques such as Fourier transform infrared spectroscopy or mass spectrometry may be sufficiently selective to provide unambiguous identification. These are often the techniques on which confirmatory methods are based.

[11] Miller, J.C., & Miller, J.N. (1993) Statistics for Analytical Chemistry, 3rd Edition, Ellis Horwood Ltd., Chichester.

Typically, a minimum of four identification points is required to meet accepted performance criteria for regulatory methods. Methods based on high resolution mass spectrometry are considered to give a higher reliability through more precise measurement of mass than can be obtained using low resolution mass spectrometry techniques. Method performance requirements for confirmatory methods based on low resolution GC/MS and LC/MS, as recently published by an international expert body[12], are given in Table 2.

Table 2: Performance requirements for relative ion intensities (sample compared to standard) using various mass spectrometric analytical techniques[9]

Relative ion intensity (% of base peak)	**GC-MS (EI) (relative)**	**GC-MS (CI), GC-MS/MS LC-MS, LC-MS/MS (relative)**
>50%	≤10%	≤20%
20% to 50%	≤15%	≤25%
10% to 20%	≤20%	≤30%

It is considered that one identification point should be assigned to each structurally significant ion fragment detected using a low resolution mass spectrometric method. When a tandem low resolution instrument, such as a “triple quadrupole” mass spectrometer is used, secondary fragments are detected from a primary fragment that is isolated in the first stage of the spectrometer. The fact that these structurally significant fragments are produced from the fragmentation of a major fragment (parent or precursor ion) associated with the molecule provides greater confidence and each such daughter or product ion is assigned a value of 1.5 identification points. A combination of a precursor ion and two product ions provides the 4 required identification points when low resolution MS/MS instruments are used in a confirmatory method.

Additional confidence is provided when high resolution mass spectrometers are used in a confirmatory method, as the high resolution provides more precise identification of the mass and may be used to predict the elemental composition of each fragment. For a single high resolution mass spectrometer, each structurally significant fragment detected is assigned a value of two identification points, while product ions generated in high resolution MS/MS experiments are assigned an identification point value of 2.5 each. In addition, at least one ion ratio must also be measured to eliminate the potential for fragments of the same mass arising from isobaric compounds of similar structure.

Other techniques, when they are used in combination, may be capable of achieving a comparable degree of selectivity as confirmatory techniques. For example, identification may be verified by combinations of methods such as:

(a) Thin layer chromatography;

(b) Element-specific gas-liquid chromatography and accompanying detection systems;

(c) Formation of characteristic derivatives followed by additional chromatography; or

(d) Determining compound specific relative retention times using several chromatographic systems of differing polarity.

Such procedures must be applicable at the designated MRLVD of the analyte. When a confirmatory method such as mass spectrometry is not available, information on the selectivity associated with the analysis of a particular veterinary drug residue in a sample may be developed from various sources[13]. This information may be captured in a structured logging document of all the information that leads to the conclusion a method has detected a particular compound in a sample, at a measured concentration as reported. While no single measurement or analysis may provide the unequivocal proof of compound identity and/or quantity present that is desired, the combined information that has been compiled provides evidence that the analyst has made a conscientious effort to arrive at a logical result consistent with the data and other information available. Examples of analytical techniques which may be suitable to meet criteria for confirmatory analytical methods are summarized in Table 3.

[12] Bethem, R., Boison, J.O., Gale, J., Heller, D., Lehotay, S., Loo, J., Musser, S., Price, P., and Stein, S. (2003) Establishing the Fitness for Purpose of Mass Spectrometric methods. *Journal of the American Society for Mass Spectrometry* **14:** 528-541.

[13] Stephany, R.W. (2003). SPECLOG – The Specificity Log. CRD-9, Codex Committee on Residues of Veterinary Drugs in Foods, 14th Session, Arlington, VA., U.S.A., March 4-7.

Table 3. Examples of detection methods suitable for the confirmatory analysis of substances, as recommended by the Miskolc Consultation[9]

Detection method	***Criterion***
LC or GC and Mass Spectrometry	If sufficient number of fragment ions are monitored
LC-DAD	If the UV spectrum is characteristic
LC – fluorescence	In combination with other techniques
2-D TLC – (spectrophotometry)	In combination with other techniques
GC-ECD, NPD, FPD	Only if combined with two or more separation techniques[a]
Derivatisation	If it was not the first choice method
LC-immunogram	In combination with other techniques
LC-UV/VIS (single wavelength)	In combination with other techniques

[a] Other chromatographic systems (applying stationary and/or mobile phases of different selectivity) or other techniques.

Although confirmatory methods are generally instrumental procedures, observation of a pathologic or other morphologic change that specifically identifies exposure to a class of veterinary drugs, could potentially be a confirmatory method, if it has sufficient sensitivity and precision.

18.4 General Performance Characteristics for Methods for Use in a Regulatory Control Programme

There are some additional considerations for selection of suitable methods for use in a regulatory control programme for veterinary drug residues in foods. Methods should be rugged (robust), cost effective, relatively uncomplicated, portable, and capable of simultaneously handling a set of samples in a time effective manner. The stability of analytes must also be established.

Ruggedness testing should be conducted using the standard factorial design approach to determine any critical control points[14]. Typical factors to include in a design include variations in reagent volumes or concentrations, pH, incubation or reaction time and temperature, reagent quality, and different batch or source of a reagent or chromatographic material. Ruggedness testing of a confirmatory method may be required if the method differs significantly from the quantitative method previously validated (if the method uses different extraction or derivatisation procedures than are used in the quantitative method).

Cost-effectiveness is the use of reagents and supplies which are readily available in the required purity from local suppliers and equipment for which parts and service are also readily available. The *method efficiency* is increased when multiple samples can be analysed at the same time. This reduces the analytical time requirements per sample and usually reduces the cost per sample, as there are certain fixed costs associated with the analysis of samples, whether done singly or in larger sets. The ability of a method to accommodate multiple samples in a batch is important when large numbers of samples must be analysed in short or fixed time frames. *Portability* is the analytical method characteristic that enables it to be transferred from one location to another without loss of established analytical performance characteristics.

Analyte stability during analysis must be established for both standards and analyte in the presence of sample material, during processing through the complete analysis for all methods used in a regulatory control programme and for typical conditions of storage while a sample is awaiting analysis. The period chosen for stability during storage should cover the expected time when sample material may be stored for all required analyses, including the use of the screening, quantitative, and confirmatory methods. It is prudent to conduct the storage study for a period which extends to at least 90 days beyond the expected time for all screening, quantitative, and confirmatory analyses to be completed and the results reported in case there is a challenge and a request for re-analysis.

19. METHOD DEVELOPMENT AND VALIDATION CONSIDERATIONS FOR RESIDUE CONTROL METHODS

19.1 Selection of Appropriate Test Material for Validation

Laboratories must demonstrate that the methods in use for analysis of regulatory samples have been suitably validated. Traditionally, the multi-laboratory method validation study has been the preferred approach to provide analytical data to define method performance characteristics. However, other models have been developed which include multi-laboratory trials with smaller numbers of laboratories than are required to conduct a full collaborative study and single laboratory validation based on rigorous in-house evaluation of method performance, supported by a Quality Management System, independent audits and analysis of proficiency or reference materials, when available.

[14] Youden, W.J., & Steiner, E.H. (1975) *Statistical Manual of the Association of Official Analytical Chemists*, AOAC International, Gaithersburg, VA.

In developing and validating a residue control method, data should be derived from three types of sample material. Control test material from non-treated animals provides information about analytical background and matrix interferences. Fortified test material, containing known amounts of the analyte added to the control material, yields information about the method's ability to recover the analyte of interest under controlled conditions. Tissues should be obtained from multiple sources to cover the variations resulting from factors such as different diets, husbandry practices, sex, and breed of animals. A minimum of six different sources of material is recommended.

In some instances, known drug free sample materials may not be available for use in residue control laboratories. In these instances an equivalent sample material may be used. Equivalent sample materials may consist of either the same matrix as the test sample matrix from an unknown source, or a different matrix from a known drug free source that closely matches the sample matrix. In all cases, the residue control laboratory must demonstrate that the equivalent sample material is free from interferences for the drug and exhibits satisfactory recovery for fortified samples. Additionally, when a material is used from an unknown source for quantitative or screening methods, it is recommended that a second method be used to demonstrate that the matrix does not contain residues of the drug. It is the responsibility of the residue control laboratory to demonstrate fitness for purpose of the equivalent sample material.

Finally, analysis of biologically incurred tissue from food producing animals that have been treated with the drug provides information about biological or other interactions that may occur when analysing residue control samples.

19.2 Measurement Uncertainty

Laboratories should provide their customers on request with information on the measurement uncertainty or statement of confidence associated with the quantitative results produced by each quantitative method. Guidance on estimation of measurement uncertainty is being developed by IUPAC and has been published by other independent scientific bodies[15].

19.3 Use of Internal Standards

Residue methods are sometimes designed using internal standards for analytical control. A properly used internal standard will compensate for some of the analytical variability of an analysis, improving precision. However, an improperly used internal standard may obscure variables that are an important part of the analytical measurement. If an internal standard is used, it should be added to a sample as early as possible in the procedure, preferably to the test material before analysis begins. The internal standard must reflect the recovery of the target analyte in a uniform and predictable fashion. An internal standard that does not mirror the behaviour of the target analyte in the method will lead to significant errors in calculation of the final result. Caution must be taken in the choice of internal standards to ensure that they do not alter the percent recovery of the analyte of interest or interfere with the measurement process. It is important to know the extent and predictability of the effects of the internal standard on an analytical method. Internal standards can greatly enhance method performance when used properly.

19.4 Environmental Considerations

If residue control methods may be subjected to widely variable physical test environments, this should be taken into account in the development and validation of these methods. Addressing these issues may help improve method ruggedness. Warmer environments may require reagents to be more thermally stable, while solvents used in the analysis will have to be less volatile and test sample requirements to be more tolerant. Cooler environments may require reagents and solvents to have different physical properties, such as lower freezing point and greater solvating characteristics, to provide effective extraction of an analyte. Environmental temperatures may influence the time required to perform an analysis, as well as influencing reaction rates, gravitational separations, and colour development. These considerations may strain efforts to standardize methods for use in broadly differing environments because of the need to adapt methods to compensate for these factors. It is important when considering the physical environment in which a method will be used to remember that volumetric glassware and many analytical instruments are calibrated to be used at specific temperatures, or within a controlled range of temperature. Operation outside these temperatures may compromise test results.

19.5 Choice of Validation Model

An analytical method developed and used in only one laboratory may have limited use in a residue control programme unless care is taken to meet the rigorous expectations for single laboratory method validation associated with accreditation under ISO/IEC-17025 or equivalent accreditation procedures for testing

---

[15] EURACHEM/CITAC Guide to Quantifying Measurement Uncertainty in Analytical Measurement, http://www.measurementuncertainty.org/mu/guide/index.html, accessed 18 September, 2007.

laboratories. The reliability of reported values may be a concern even though strong quality control procedures may have been employed, unless supported by data from an on-going proficiency programme, comparison with a suitable method validated in an inter-laboratory trial or other forms of inter-laboratory comparison of results. Ideally, a method should be validated by at least three laboratories. Methods which have been carefully validated in a single laboratory with inclusion of properly designed ruggedness tests should be able to successfully undergo a collaborative study involving at least eight different laboratories.

The principles for conducting a single laboratory method validation, a multi-laboratory method trial or a collaborative study of a residue control method are the same. Samples for evaluating method performance should be unknown to the analyst, in randomised replicates, containing the residue near the MRLVD or other target concentration, as well as samples with the analyte above and below the concentration of interest, and test material blanks. A minimum of three individual datasets should be generated over three analysis periods, on at least three separate occasions (at least one day apart), preferably with replicate analysis, to improve statistical evaluation of method performance and provide an estimate of inter-day variability. It should be noted that these are only minimal requirements. The establishment of statistically-based performance standards for methods is enhanced by increasing the number of independent analysts and laboratories testing the method, as well as by the number of samples tested. In a single-laboratory validation, it is recommended that the method should be tested by multiple analysts to provide appropriate measures of within-laboratory performance. Expanding the validation to include other laboratories, preferably to the number required for a collaborative study, is recommended. Analyses of blind duplicates, as required in the collaborative study protocol[7] in only eight laboratories, with one or two animal species and tissues, yields limited quality estimates for overall repeatability and reproducibility. The validation of a collaboratively studied method can be extended to include additional tissues and species in a subsequent study conducted by a single expert laboratory, as required.

19.6 Quality Management Systems

A Quality Management System is an essential component of residue analysis. It both monitors those factors associated with the analysis of a sample by an analyst and provides the oversight by independent reviewers to ensure that the analytical programme is performing in an acceptable manner. The use of an accredited Quality Management System is invaluable to support decision-making for residue control agencies, improving the reliability of analytical results, and providing quality data for residue control programmes to demonstrate food safety to consumers, producers, and law making bodies regarding residues of veterinary drugs in food. The establishment of quality measures consistent with the principles published by IUPAC is recommended for regulatory control laboratories.

APPENDIX A

SAMPLING STRATEGIES

A1. NON-BIASED SAMPLING

A1.1 Purpose

Non-biased sampling is designed to provide profile information, especially as to the extent of application or performance of a control or assurance system for a specified animal/food population over a defined period.

A1.2 Statistical considerations on sampling population size

The number of samples for non-biased sampling protocols should be statistically based and may be influenced by the size of the population (where less than 5000), the prevalence of non-compliance determined to be significant, the confidence to be placed in the results as well as economic considerations.

The number of samples based on the binomial distribution will always be equal to or greater than the required number of samples based on the hypergeometric distribution[16].

If the size of the population is small the effect of sampling without replacement is significant and the sampling distribution should be based on the hypergeometric distribution.

In populations larger than 5000 units the effect of sampling without replacement is negligible. Thus the binomial distribution can be used to determine an appropriate number of samples.

The number of samples for a defined confidence will be effectively constant for populations exceeding 5000 units.

A1.3 Sampling Confidence reporting

Where non-compliant results are detected it is possible to derive a crude estimate of the likely prevalence in the general population.

However, where no non-compliant results are found then any statements about prevalence need to be stated with a defined confidence that the prevalence of non-compliant results does not exceed a specified percentage.

The number of samples required to give a required statistical assurance can be read from Table 4. Other scientifically based statistical protocols may also be used.

Table 4: Number of samples required to detect at least one non-compliant result with pre-defined probabilities (90, 95, and 99 percent) in a population having a known non-compliance prevalence.

Non-compliant prevalence (% in a population)	Minimum number of samples required to detect a non-compliant result with a confidence level of:		
	90%	95%	99%
35	6	7	11
30	7	9	13
25	9	11	17
20	11	14	21
15	15	19	29
10	22	29	44
5	45	59	90
1	230	299	459
0.5	460	598	919
0.1	2302	2995	4603

The probability of failing to detect a specified prevalence of non-compliant results associated with a specified targeting mechanism can be read off Table 5 below. Because of the low efficacy of sampling protocols to detect low prevalences of non-compliance, other assurance mechanisms are more important where a low prevalence of non-compliance is expected.

[16] In the probability theory and statistics, the *hypergeometric distribution* is a discrete (consisting of unconnected distinct parts) probability distribution that describes the number of successes in a sequence of *n* draws from a finite population without replacement.

Table 5: Probability of failing to detect a non-compliance

Prevalence (%)	Number of animals/units of product in sample tested									
	5	**10**	**25**	**50**	**75**	**100**	**200**	**250**	**500**	**1000**
1	0.951	0.904	0.779	0.605	0.471	0.366	0.134	0.081	0.007	0.000
2	0.904	0.817	0.603	0.364	0.220	0.133	0.018	0.006	0.000	
3	0.859	0.737	0.467	0.218	0.102	0.048	0.002	0.000		
4	0.815	0.665	0.360	0.130	0.047	0.017	0.000			
5	0.774	0.599	0.277	0.077	0.021	0.006				
6	0.734	0.539	0.213	0.045	0.010	0.002				
7	0.696	0.484	0.163	0.027	0.004	0.001				
8	0.659	0.434	0.124	0.015	0.002	0.000				
9	0.590	0.389	0.095	0.009	0.001					
10	0.528	0.349	0.072	0.005	0.000					
12	0.470	0.279	0.041	0.002						
14	0.418	0.221	0.023	0.001						
16	0.371	0.175	0.013	0.000						
18	0.328	0.137	0.007							
20	0.254	0.107	0.004							
24	0.193	0.064	0.001							
28	0.193	0.037	0.000							
32	0.145	0.021								
36	0.107	0.012								
40	0.078	0.006								
50	0.031	0.001								
60	0.010	0.000								

A2. DIRECTED OR TARGETED SAMPLING

A2.1 Purpose

Directed or targeted sampling protocols are designed to place a greater intensity of inspection/audit on suppliers or product considered to possibly have a greater potential than the general population of being non-compliant.

It is not possible to extrapolate from non-compliant results to draw conclusions about the general population because a sub-population which is considered to have greater chance of non-compliance is being sampled (biased sampling).

However, if compliant results confirm non-biased programme results, they provide increased assurance that the system is working effectively.

APPENDIX B

SAMPLING OF COMMODITIES

B1. SCOPE

This Appendix applies to the following commodities: primary food commodities of animal origin and processed products of animal origin made from primary food appearing in Tables A, B and C of this appendix, and honey of the following origins and/or processing method:

(a) Blossom or nectar honey that comes mainly from nectaries of flowers;

(b) Honeydew honey that comes mainly from secretions of or on living parts of plants;

(c) Comb honey stored by bees in the cells of freshly built broodless combs, and sold in sealed whole combs or sections of such combs;

(d) Extracted honey obtained by centrifuging decapped broodless combs;

(e) Pressed honey obtained by pressing broodless combs with or without the application of moderate heat.

B2. DEFINITIONS

Lot means an identifiable group of animals or quantity of animal product intended for food use and determined to have common characteristics, such as origin variety, type of packing, packer or consignor, or markings, by the sampling official. Several lots may make up a consignment.

Consignments means an identifiable group of animals or quantity of animal product intended for food use as described on a particular contractor's shipping document. Lots in a consignment may have different origins or may be delivered at different times.

Primary sample means a quantity of representative biological material taken from a single animal (or group of animals) or from one place in the lot. When the quantity is inadequate for residue analysis, samples from more than one animal (or group of animals) or more than one location in the lot can be combined for the primary sample (such as poultry organs).

Bulk sample means the combined total of all the primary samples taken from the same lot.

Final laboratory sample means the primary or bulk sample, or a representative portion of the primary or bulk sample, intended for laboratory analysis.

Final laboratory test portion means the representative portion of the final laboratory sample on which an analysis is conducted. The entire laboratory sample may be used for analysis in some cases but typically will be sub-divided into representative test portions for analysis. It is prepared by combining and thoroughly mixing the primary samples.

Lot of honey means a discrete quantity of honey delivered for distribution at one time, and determined to have common characteristics, such as origin, variety, type of packing, packer or consignor, or markings, by the sampling official.

Consignment of honey means discrete quantity of honey as described on a particular contractor's shipping document. A consignment may be made up of different lots.

Primary honey sample means a quantity of honey taken from one place in the lot, unless this quantity is inadequate for the residue analysis. When the quantity is inadequate, samples from more than one location can be combined for the primary sample.

B3. SAMPLING PROCEDURES

Samples must be collected by those officially authorized for this purpose.

Each lot to be examined must be sampled separately.

During collection and processing care must be taken to prevent contamination or other changes in the samples which would alter the residue, affect the analytical determination, or make the laboratory test portion not representative of the bulk or laboratory sample.

Guidance on sample type and quantity for different commodities is provided in Table A: Meat and Poultry Products; Table B: Milk, Eggs and Dairy Products and Table C: Aquaculture Products. The following are general instructions:

(a) Each primary sample should be taken from a single animal (or group of animals) or unit in a lot, and when possible, be selected randomly;

(b) When several animals are required for adequate sample size of the primary sample (e.g. poultry liver), the samples should be collected consecutively after initial random selection;

(c) Frozen product should not be thawed before sampling;

(d) Canned or packaged product should not be opened for sampling unless the unit size is at least twice the amount required for the final laboratory sample. The final laboratory sample should contain a representative portion of juices surrounding the product;

(e) Unopened cans or packages which constitute a final laboratory sample should be sent unopened and intact to the laboratory for analysis;

(f) The contents of cans or packages opened by the authorised inspector should be frozen as described in paragraph 170d before dispatch to the laboratory for analysis;

(g) Large, bone-containing units of product (i.e. prime cuts) should be sampled by collecting edible product only as the primary sample;

(h) When portions of single unit are less than described as a primary sample, additional sample units need to be taken to satisfy bulk sample requirements.

(i) Portions remaining of final laboratory samples should be frozen and stored in conditions which will maintain the sample integrity.

The number of primary samples collected will depend on if a lot is considered suspect.

A lot is suspect if there is:

(a) A history of non-compliance with the MRLVD;

(b) Evidence of contamination during transport;

(c) Signs of toxicosis (systemic poisoning) observed during ante- or post-mortem inspection; or

(d) Other relevant information available to the authorised inspection official.

A minimum of six to a maximum of thirty primary samples should be collected from a suspect lot. When the suspected residues are is expected to occur throughout the lot the smaller number of samples is sufficient.

Imports from countries that do not run verification programmes for compliance with MRLVDs should be sampled as suspect lots.

B4. SPECIFIC SAMPLE PREPARATION INSTRUCTIONS FOR HONEY

(a) Collect 250 mL of liquid or strained honey after the following preparations as applicable;

(b) Liquidise comb honey: Cut across top of comb, if sealed, and separate completely from comb by straining through a sieve the meshes of which are made by so weaving wire as to form square opening of 0.500 mm by 0.500 mm (ISO 565-1990)[17].

(c) If foreign matter, such as wax, sticks, bees, particles of comb, etc., is present, heat sample to 40°C in water bath and strain through cheesecloth in hot-water-funnel before sampling.

When a sample is free from granulation mix thoroughly by stirring or shaking; if granulated, place closed container in water-bath without submerging, and heat for 30 min at 60°C; then if necessary heat at 65°C until liquefied. Occasional shaking is essential. Mix thoroughly and cool rapidly as soon as the sample liquefies.

B5. STATISTICAL CONCERNS

For non-suspect lots a statistically-based, non-biased sampling programme is recommended. Any of the following types of sampling can be used.

B5.1 Stratified random sampling

Where consignments are commingled simple random criteria cannot be applied and stratified random sampling should be considered.

In stratified random sampling the consignment is divided into non-overlapping groups or strata, e.g. geographical origin, genders, time. A sample is taken from each stratum.

[17] Such sieve could be replaced by US sieve with No. 40 standard screen (size of opening 0.420 mm).

Homogeneity within each stratum is better than in the whole population. Countries or geographic regions are considered natural strata based on uniformity in agricultural practices.

Time strata (e.g., month, quarter) are commonly used for convenience, efficiency, and detection of seasonal variability. Random number tables[18] or other objective techniques should be used to ensure that all elements of a population have an equal and independent chance of being included in the sample.

B5.2 Systematic sampling

In systematic sampling units are selected from the population at a regular interval *(*e.g., once an hour, every other lot, etc.).

It may be applied when there is reliable information on product volumes to determine the sampling interval that will provide the desired number of samples over time. However:

(a) If the sampling system is too predictable, it may be abused;

(b) Consignments need to be homogeneous, because systematic sample units are uniformly distributed over the population.

B5.3 Biased or estimated worst case sampling

In biased or estimated worst case sampling, investigators use their judgement and experience regarding the population, lot, or sampling frame to decide which primary samples to select.

The population group anticipated to be at greatest risk may be identified, but no general conclusion should be made about the population sampled from the data collected (non-random samples).

B6. PREPARATION OF LABORATORY SAMPLES

The final laboratory sample is sent for analysis.

Some national/regional legislation/regulation may require that the final laboratory sample is sub-divided into two or more portions for separate analyses. Each portion should be representative of the final laboratory sample. Precautions indicated under *sampling procedures* should be observed.

The laboratory test portion should be prepared from the final laboratory sample by an appropriate method of reduction.

B7. SHIPMENT OF LABORATORY SAMPLES

Final laboratory samples should be prepared as follows:

(a) Each sample should be placed in a clean, thermally insulating, chemically inert container to protect the sample from contamination, defrosting and damage in shipping;

(b) The container should be sealed so that unauthorized opening is detectable;

(c) The container should be sent to the laboratory as soon as possible, after taking precautions against leakage and spoilage;

(d) For shipping, all perishable samples should be frozen to minus 20°C, immediately after collection, and packed in a suitable container that retards thawing. Freezer packs or other suitable refrigerants should be used to maintain freezer temperatures during shipment. Samples and freezer packs should be fully frozen to minus 20°C prior to dispatch;

(e) Replicate portions of the final laboratory sample which may be retained as required by national/regional legislation or as an administrative policy should be placed in a clean, chemically inert container to protect the sample from contamination, sealed so that unauthorized opening is detectable and stored under suitable conditions to prevent a change in the product or any residues it may contain in case future analysis is required for comparison with analytical results obtained on the sample material submitted to the laboratory.

B8. RESULT INTERPRETATION IN THE LABORATORY

For purposes of control, the MRLVD is applied to the residue concentration found in each laboratory sample taken from a lot.

---

[18] Random number tables consist of a randomly generated series of digits (0-9). To improve readability there are spaces between every e.g. every 4th digit and between every 10th rows. Reading can begin anywhere (at random) but having started has to continue across the line or down a column and NOT jump about. Example: extract from a table of random sampling numbers: 3680 2231 8846 5418 0498 5245 7071 2597.

Lot compliance with a MRLVD is achieved when the mean result for analysis of the laboratory test portions does not indicate the presence of a residue which exceeds the MRLVD.

B9. SAMPLING RECORDS

Each primary or bulk sample and each final laboratory sample should be uniquely linked to a record with the type of sample, analyses required, its origin (e.g., country, state, or town), its location of collection, date of sampling, and additional information required for follow-up action if necessary.

If there is a deviation from recommended sampling procedures, records accompanying the sample should describe procedures actually followed in detail.

GUIDANCE ON SAMPLE TYPE AND QUANTITY FOR DIFFERENT COMMODITIES

Table A: Meat and poultry products

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
I. Group 030 (Mammalian Meats)		
A. Whole carcass or side, unit weight normally 10 kg or more	Collect diaphragm muscle, supplement with cervical muscle, if necessary, from one animal.	500 g
B. Small carcass (e.g. rabbit)		500 g after removal of skin and bone
C. Fresh/chilled parts		
1. Unit minimum weight of 0.5 kg, excluding bone (e.g. quarters, shoulders, roasts)	Collect muscle from one unit.	500 g
2. Unit weighing less than 0.5 kg (e.g. chops, fillets)	Collect the number of units from selected container to meet laboratory sample size requirements.	500 g after removal of bone
D. Bulk frozen parts	Collect a frozen cross-section from selected container, or take muscle from one large part.	500 g
E. Retail packaged frozen/chilled parts, or individually wrapped units for wholesale	For large cuts, collect muscle from one unit or take sample from number of units to meet laboratory sample size requirements.	500 g after removal of bone
Ia. Group 030 (Mammalian Meats where MRL is expressed in carcass fat)		
A. Animals sampled at slaughter	See instructions under II. Group 031.	
B. Other meat parts	Collect 500 g of visible fat, or sufficient product to yield 50-100 g of fat for analysis. (Normally 1.5-2.0 kg of product is required for cuts without trimmable fat).	Sufficient to yield 50-100 g of fat
II. Group 031 (Mammalian Fats)		
A. Large animals sampled at slaughter, usually weighing at least 10 kg	Collect kidney, abdominal, or subcutaneous fat from one animal.	500 g
B. Small animals sampled at slaughter[(a)]	Collect abdominal and subcutaneous fat from one or more animals.	500 g
C. Bulk fat tissue	Collect equal size portions from 3 locations in container.	500 g
III. Group 032 (Mammalian Edible Offal)		
A. Liver	Collect whole liver(s) or portion sufficient to meet laboratory sample size requirements.	400 - 500 g
B. Kidney	Collect one or both kidneys, or kidneys from more than one animal, sufficient to meet laboratory sample size requirement. Do not collect from more than one animal if size meets the low range for sample size.	250 - 500 g
C. Heart	Collect whole heart or ventricle portion sufficient to meet laboratory sample size requirement.	400 - 500 g

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
D. Other fresh/chilled or frozen, edible offal product	Collect portion derived from one animal unless product from more than one animal is required to meet laboratory sample size requirement. A cross-section can be taken from bulk frozen product.	500 g
IV. Group 036 (Poultry Meats)		
A. Whole carcass of large bird, typically weighing 2-3 kg or more (e.g. turkey, mature chicken, goose, duck)	Collect thigh, leg, and other dark meat from one bird.	500 g after removal of skin and bone
B. Whole carcass of bird typically weighing between 0.5-2.0 kg (e.g. young chicken, duckling, guinea fowl)	Collect thigh, legs, and other dark meat from 3-6 birds, depending on size.	500 g after removal of skin and bone
C. Whole carcasses of very small birds typically weighing less than 500 g (e.g. quail, pigeon)	Collect at least 6 whole carcasses	250 - 500 g of muscle tissue
D. Fresh/chilled or frozen parts		
1. Wholesale package		
a. Large parts	Collect an interior unit from a selected container.	500 g after removal of skin and bone
b. Small parts	Collect sufficient parts from a selected layer in the container	500 g after removal of skin and bone
2. Retail packaged	Collect a number of units from selected container to meet laboratory sample size requirement.	500 g after removal of skin and bone
IVa. Group 036 (Poultry Meats where MRLVD is expressed in carcass fat)		
A. Birds sampled at slaughter	See instructions under V. Group 037	
B. Other poultry meat	Collect 500 g of fat or sufficient product to yield 50-100 g of fat. (Normally, 1.5-2.0 kg is required.)	500 g of fat or enough tissue to yield 50-100 g of fat
V. Group 037 (Poultry Fats)		
A. Birds sampled at slaughter	Collect abdominal fat from 3-6 birds, depending on size.	Sufficient to yield 50-100 g of fat
B. Bulk fat tissue	Collect equal size portions from 3 locations in container.	500 g
VI. Group 038 (Poultry Edible Offal)		
A. Liver	Collect 6 whole livers or a sufficient number to meet laboratory sample requirement.	250 - 500 g
B. Other fresh/chilled or frozen edible offal product	Collect appropriate parts from 6 birds. If bulk frozen, take a cross-section from container.	250 - 500 g

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
VII. Class E - Type 16 (Secondary Meat and Poultry Products)		
A. Fresh/chilled or frozen comminuted product of single species origin	Collect a representative fresh or frozen cross-section from selected container or packaged unit.	500 g
B. Group 080(Dried Meat Products)	Collect a number of packaged units in a selected container sufficient to meet laboratory sample size requirements.	500 g, unless fat content is less than 5% and MRLVD is expressed on a fat basis. Then 1.5-2.0 kg is required.
VIII. Class E-Type 18 (Manufactured, single ingredient product of animal origin)		
A. Canned product (e.g. ham, beef, chicken), unit size of 1 kg or more	Collect one can from a lot. When unit size is large (greater than 2 kg), a representative sample including juices may be taken.	500 g, unless fat content is less than 5% and MRLVD is expressed on a fat basis. Then 1.5-2.0 kg is required.
B. Cured, smoked, or cooked product (e.g. bacon slab, ham, turkey, cooked beef), unit size of at least 1 kg	Collect portion from a large unit (greater than 2 kg), or take whole unit, depending on size.	500 g, unless fat content is less than 5% and MRLVD is expressed on a fat basis. Then 1.5-2.0 kg is required.
IX. Class E - Type 19 (Manufactured, multiple ingredient, product of animal origin)		
A. Sausage and luncheon meat rolls with a unit size of at least 1 kg	Collect cross-section portion from a large unit (greater than 2 kg), or whole unit, depending on size.	500 g

(a) When adhering fat is insufficient to provide a suitable sample, the sole commodity without bone, is analysed and the MRL will apply to the sole commodity.

Table B: Milk, eggs, dairy products

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
I. Group 033 (Milks)		
Whole liquid milk raw, pasteurised, UHT & sterilized	In bulk. Mix thoroughly and immediately take a sample by means of a dipper. In retail containers. Take sufficient units to meet laboratory sample size requirements.	500 mL
II. Group 082 (Secondary Milk Products)		
A. Skimmed milk - skimmed and Semi-skimmed	As for whole liquid milk Bulk containers (barrels, drums). Mix the contents carefully and scrape adhering material from the sides and bottom of the container. Remove 2 to 3 litres, repeat the stirring and take a 500 mL sample.	500 mL
B. Evaporated milk - evaporated full cream & skimmed milk	Small retail containers. Take sufficient units to meet laboratory sample size requirements.	500 mL

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
C. Milk powders		
1. Whole	Bulk containers. Pass a dry borer tube steadily through the powder at an even rate of penetration. Remove sufficient bores to make up a sample of 500 g. Small retail containers. Take sufficient units to meet laboratory sample size requirements..	500 g
2. Low fat	As for whole milk powders	500 g
III. Group 087 (Derived Milk Products)		
A. Cream - fresh, frozen & UHT; single, whipping, whipped, double & clotted	Bulk containers. Plunge to ensure thorough mixing moving the plunger from place to place avoiding foaming, whipping and churning. Take a 200 ml sample by means of a dipper. Small containers. Take sufficient units to meet laboratory sample size requirements.	200 mL
B. Butter - including whey butter and low fat spreads containing butterfat	In bulk. Take two cores or more of butter so that the minimum total sample weight is not less than 200 g In pats or rolls. For units weighing over 250 g divide into four and take opposite quarters. For units weighing less than 250 g take one unit as sample.	200 g
C. Butter oil - including anhydrous butte roil and anhydrous milk fat	Mix thoroughly and take a 200 g sample.	200 g
IV. Group 090 (Manufactured Milk Products - single ingredient)		
A. Yoghurt - natural, low fat through to full cream	Select number of units sufficient to meet laboratory requirements.	500 g
B. Cheeses - all varieties	Make two cuts radiating from the centre of the cheese if the cheese has a circular base, or parallel to the sides if the base is rectangular. The piece removed should meet the laboratory sample size requirements. For small cheeses and wrapped portions of cheese take sufficient units to meet laboratory sample requirements.	200 g
V. Group 092 (Manufactured Milk Products - multi-ingredient)		
A. Dairy ice cream - only ice cream containing 5% or greater of milk fat	Select block or units sufficient to meet laboratory sample size requirements.	500 mL
B. Processed cheese preparations	Select units sufficient to meet laboratory sample size requirements.	200 g
C. Flavoured yoghurt	As for natural yoghurt.	500 g
D. Sweetened condensed Milk	As for evaporated milk.	500 mL

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
VI. Group 039 (Eggs and Egg Products)		
A. Liquid and frozen eggs	Use sample schedule. Sub sample size will be 250 mL liquid or 500 mL packed shavings from aseptic drillings into containers.	500 g
B. Dried egg products	Use sample schedule. For containers of 500 g or less or 25 mL or less, collect a minimum of 2 units per sub sample. For containers of 500 g to 10 kg select 1 unit per sub sample. For containers of 10 kg or more collect 1 kg from each unit sampled. Collect with aseptic technique.	500 g
C. Shell eggs		
1. Retail packages	Use sample schedule. Sub sample size is 12 eggs.	500 g or 10 whole eggs
2. Commercial cases	For 15 cases or less collect 12 eggs from each case, minimum of 24 eggs. For 16 or more cases collect 12 eggs from 15 random cases.	500 g or 10 whole eggs

Table C: Aquaculture products

COMMODITY	INSTRUCTIONS FOR COLLECTION	MINIMUM QUANTITY REQUIRED FOR LABORATORY SAMPLE
VII. Class B – Type 08 (Aquatic Animal Products)		
A. Packaged fish – fresh, frozen, smoked, cured, or shellfish (except oysters)		
1. Bulk package	Collect sufficient units from a selected package to meet laboratory sample size.	500 g of edible tissue
2. Retail package	Collect sufficient units from selected packages to meet laboratory sample size.	500 g of edible tissue
B. Bulk fish	Collect edible tissue from sufficient fish, depending on size.	500 g of edible tissue
C. Bulk Shellfish	Collect sufficient shellfish, depending on size.	500 g of edible tissue
VII. Class E – Type 17 (Derived Edible Products of Aquatic Animal Origin)		
A. Canned fish and shellfish products (except oysters)	Collect sufficient tissue to meet laboratory sample size.	500 g of edible tissue
B. Other fish and shellfish products	Use sample schedule. Collect primary samples to meet laboratory sample size	500 g

APPENDIX C

PERFORMANCE CHARACTERISTICS FOR MULTI-RESIDUE METHODS (MRMs) FOR VETERINARY DRUGS

C1. SCOPE

The purpose of this Appendix is to describe the performance characteristics/parameters that a multi-residue method (MRM) should have in order to provide internationally acceptable confidence in the method to produce results suitable for evaluating the residues of veterinary drugs for either domestic programmes or in international trade. The uses may include screening, quantification, and/or confirmation, each having different performance requirements.

This Appendix is applicable to MRMs used to analyse all residues of veterinary drugs and substances which may be used as veterinary drugs. These MRMs include certain pesticides which have veterinary uses and which may be present as residues in commodities. Guidance on the validation of multi-residue methods for non-veterinary use of pesticides is contained in CAC/GL 40-1993: *Guidelines on good laboratory practice in residue analysis*.

In this Appendix, a MRM is considered to be a method which includes three or more analytes in the same class or more than one class of veterinary drugs in its scope. These MRMs may be used for screening samples for the possible presence of veterinary drugs or quantitative and/or confirmatory analyses. This guidance covers all three types of situations. It should be noted that a validated MRM may include some analytes where performance characteristics for quantitative analysis have been fully validated and other analytes where precision and/or recovery criteria for quantitative analysis or the data requirements for confirmation of the residue are not available. However, those analytes should be clearly identified in the method and must not be used for those purposes until they have been validated and/or demonstrated to be fit for purpose.

C2. DEFINITIONS

Compliant or negative result: A result indicating that the analyte is not present at or above the lowest calibrated concentration. (see also Limit of Detection in CAC/GL 72-2009).

Confirmatory method: A method that provides complete or complementary information enabling the analyte to be identified with an acceptable degree of certainty at the concentration of interest.

Decision Limit ($CC\alpha$): Limit at which it can be decided that the concentration of the analyte present in a sample truly exceeds that limit with an error probability of α (false positive).

Detection capability ($CC\beta$): Smallest true concentration of the analyte that may be detected, identified and quantified in a sample with an error probability of ß (false negative).

Incurred residue: Residue of an analyte in a matrix arising by the route through which the trace concentrations would normally be expected by treatment or dosing according to intended use, as opposed to residues from laboratory fortification of samples.

Matrix: Material or component sampled for analytical studies, excluding the analyte.

Matrix blank: Sample material containing no detectable concentration of the analytes of interest.

Method: The series of procedures from receipt of a sample for analysis through to the production of the final result.

Multi-residue method (MRM): Method which is suitable for the screening, confirmation and quantification of a range of analytes, usually in a number of different matrices and includes three or more analytes in the same class or more than one class of veterinary drugs in its scope.

Presumptive positive or suspect result: A result suggesting the presence of the analyte with a concentration at or above the lowest calibrated concentration.

Positive result: A result indicating that the analyte has been confirmed to be present at or above the lowest calibrated concentration.

Quantitative method: A method capable of producing results, expressed as numerical values in appropriate units, with accuracy and precision which are fit for the purpose. The degree of precision and trueness must comply with the criteria specified in Table 1 of the main text.

Sample preparation: Procedure used, if required, to convert the laboratory sample into the analytical sample by removal of parts not to be included in the analysis.

Sample processing: The procedure(s) (e.g. cutting, grinding, mixing) used to make the analytical sample acceptably homogeneous with respect to the analyte distribution prior to removal of the analytical portion.

Screening method: A method used to detect the presence of an analyte or class of analytes at or above the minimum concentration of interest.

C3. SUMMARY OF PERFORMANCE PARAMETERS TO BE CHARACTERISED AND DEFINED FOR MULTI-RESIDUE ANALYTICAL METHODS

The following characteristic parameters need to be measured for every analyte and for each matrix under study, as applicable:

(a) Selectivity

(i) Freedom from interferences

(ii) Matrix effects – characterised and controlled by the method if they occur**.**

(iii) Qualitative, quantitative, and/or confirmatory detector response parameters determined (and CCβ for screening analyses where this is included below to cover cut-off or threshold limits)

(b) Calibration

(i) Sensitivity

(ii) Calibration range

(iii) Calibration function

(iv) LOD and LOQ, and/or CCα and CCβ

(c) Reliability of results

(i) Recovery

(ii) Accuracy (trueness and precision)

(iii) Measurement uncertainty

(iv) Robustness (ruggedness) testing including identification of critical control points and possible stopping points

(d) Stability of Analytes

(i) Stability in sample extracts and standard solutions;

(ii) Stability under sample processing and analysis

(iii) Stability under frozen storage and freeze-thaw cycle conditions.

(e) Incurred residue studies (if suitable materials are available)

(i) Verify that incurred residues are as effectively extracted as fortified analytes

(ii) Verify performance of any steps included in method to release chemically bound residues where required.

(iii) Verify consistency of recovery and precision

C4. PERFORMANCE CHARACTERISTICS FOR MRMs

It should be understood that the performance characteristics listed in paragraph 4 should be defined and measured for every analyte listed in the scope of the fully optimised multi-residue method. This is best done after it has been determined that method development and/or modification has been completed and the analytical method is not to be subjected to any additional changes or modifications. In this regard, the concepts involved are very similar to those for determining the performance characteristics of an analyte in a single analyte method elaborated in the main text (see Section 18). To avoid repetition, only differences from single analyte consideration will be highlighted in this Appendix.

The requirement on MRMs to successfully detect residues of a variety of different veterinary drugs in a complex food matrix can be expected to result in an increased risk of interference by other material from the sample matrix compared to single analyte methods. If the MRM is required to analyse different matrices or a matrix from different species the risk is increased. This necessitates particular emphasis on performance characteristics related to detection capability and selectivity when considering the performance of MRMs.

C5. PERFORMANCE CHARACTERISTICS OF MRMS FOR SCREENING ANALYSIS

MRMs for screening analysis are usually qualitative in nature and often cover a range of analytes, species and matrices, with the objective being to differentiate samples that contain no detectable residues above a threshold or cut-off value (”negatives/compliant”) from those that may contain residues above that value (”positives/presumptive positives/suspect positives”).

Screening methods for approved veterinary drugs should demonstrate a selectivity of 90% with 95% confidence and sensitivity at the lowest concentration at which the target analyte may be reliably detected within defined statistical limits, usually 95% confidence limit. For regulatory purposes, these screening methods can tolerate a small number of “false positive” results, as any screen “positive/presumptive positive/suspect positive” sample should be carried forward for additional confirmatory and/or quantitative analysis to identify, confirm and/or quantify the presence of the “suspect” residue. For all other veterinary drugs which are NOT approved for use, this Appendix may be used to inform decisions on the performance criteria which may need to be developed.

Criteria for identifying cut-off or threshold limits for screening methods are given in the main text (see Section 18.1).

C6. PERFORMANCE CHARACTERISTICS OF MRMs FOR QUANTITATIVE ANALYSIS

The requirement to recover a range of different veterinary drug residues in one extraction increases the potential for compromised selectivity in MRMs compared to single analyte methods. Using less selective extraction and clean-up procedures is likely to result in greater co-extracted matrix material in the final extract. The nature and quantities of such co-extracted material can vary markedly depending on the history of the individual sample. Particular care is therefore required when setting criteria for the precision and trueness of MRMs to ensure that quantification will not be affected by interference from other compounds present in the sample matrix. It is recommended that MRMs used to support Codex MRLs should meet the performance standards for trueness and precision listed in Table 1 of the main text. To ensure that the effects of different samples are taken into account when assessing performance against these criteria, it is recommended that determinations of these parameters follow the guidance in the main text (see Section 18.2). The intermediate precision for recovery of analytes fortified into these different samples should be used for comparison to the criteria in Table 1 of the main text rather than the repeatability precision.

However, where no guidance is available to provide a target concentration for a specific analyte, a value based on an assessment of public health risk, and not based on the detection limits of the available analytical instrumentation may be considered.

It is becoming increasingly common in analytical methods for veterinary drug residues in foods to base the quantitative determination on a standard curve prepared by addition of standard to known blank representative matrix material prior to analyte extraction at a range of appropriate concentrations that bracket the target concentration. Use of such a method matrix-matched standard curve for calibration inherently incorporates a recovery correction into the analytical results obtained but may introduce a new bias related to the behaviour of the particular blank matrix used to construct the standard curve. It is recommended that the trueness of methods that employ matrix-matched calibration curves follow the guidelines provided in the main text (see Section 18.2).

Alternative approaches may be applied to method validation that use the parameters Decision Limit (CCα) and Detection Capability (CCß). These two parameters incorporate a consideration of measurement uncertainty.

C7. PERFORMANCE CHARACTERISTICS FOR MRMS FOR CONFIRMATORY METHODS

The necessary steps to positive identification are for the expert judgement of the analyst and particular attention should be paid to the choice of a method that would minimise the effect of interfering analytes. Ultimately, it is the responsibility of the analyst to make choices, provide supporting data, and interpret results according to scientific principles and qualified judgement as outlined in the main text (see Section 18.3.

Method performance requirements for confirmatory methods based on low resolution gas chromatography mass spectrometry (GC-MS) and liquid chromatography mass spectrometry (LC-MS) listed in Table 2 of the main text have been extended to include situations where the relative ion intensity may be less than 10%. Under these conditions, a 50% relative ion intensity between standard and sample is acceptable.

Table 1 in this Appendix lists the number of identification points (IPs) earned for a combination of mass spectrometry based analytical techniques and provides necessary and sufficient criteria for confirmatory analysis. Typically, a minimum of four identification points is required to meet accepted performance criteria

for regulatory methods. Therefore, a combination of a precursor ion and two product ions will provide the four IPs required when low resolution MS/MS instruments are used in a confirmatory method. Examples of non-MS based detection methods are listed in Table 3 in the main text.

Regardless of the mass spectrometer resolution, at least one ion ratio must also be measured to eliminate the potential for fragments of the same mass arising from isobaric compounds of similar structure. Retention times, or better still relative retention times, should also be determined to avoid the potential for false identifications when using mass spectrometers for detection.

Non-magnetic sector type high-resolution mass spectrometers (HRMS) are becomingly increasingly more affordable and commonly used. If using this equipment, it is suggested that confirmation of a compound be based on the high mass accuracy and the resolving power of the mass spectrometer**.**

C8. VALIDATION OF THE FULLY CHARACTERIZED MRM

Determination of the parameters in paragraph 4 for all the analytes and matrices listed in the scope of a MRM will allow an objective assessment to be made of the fitness-for-purpose of the analytical method for use in a regulatory control programme. For screening methods, analytes whose measured performance parameters in a set of validation experiments are achieved in ≥ 90% of the measurements taken at each analyte/matrix/concentration combination could be considered acceptable for inclusion in the method.

Section 19.1 recommends the use of biologically incurred material in the characterisation and validation of analytical methods where possible, but the cost of generating such incurred material for the validation of each analyte in a MRM could be prohibitive. However, where it is economically feasible and possible to administer several different veterinary drugs to a food animal, incurred material may be generated for a few carefully selected analytes representative of drug classes and/or groups based on their prevalence of use and potential for causing residues that exceed established MRLs. The target incurred concentration should be close to the MRL or expected concentration.

Alternative protocols may be used for validation of MRMs, adapted as necessary for individual circumstances.

Table 1: Examples of the number of identification points (IPs) earned for a range of mass spectrometric detection techniques and combinations thereof (n = an integer)

Technique	Source of Identification	Number of Identification Points (IPs)
GC-MS (EI[a] or CI[b])	n characteristic ions	N
GC-MS (EI +CI)	2 (EI) + 2 (CI)	4
GC-EIMS or GC-CIMS (2 derivatives)	2 (Derivative A) + 2 (Derivative B)	4
LC-MS	n characteristic ions	N
GC-MS/MS[c]	1 precursor ion + 2 product ions	4
LC-MS/MS[d]	1 precursor ion + 2 product ions	4
GC-MS/MS	2 precursor ions, each with 1 product ion	5
LC-MS/MS	2 precursor ions, each with 1 product ion	5
LC-MS/MS/MS	1 precursor, 1 product ion and 2 2nd generation product ions	5.5
HRMS	N	2n
GC-MS and LC-MS	2 + 2	4
GC-MS and HRMS	2 + 1	4
LC-HRMS/MS and GC-HRMS/MS	1 precursor ion + 2 product ions	6

[a] Electron ionisation (EI)

[b] Chemical ionisation (CI)

[c] Gas chromatography tandem mass spectrometry (GC-MS/MS)

[d] Liquid chromatography tandem mass spectrometry (LC-MS/MS)