CODEX ALIMENTARIUS COMMISSION

Food and Agriculture Organization of the United Nations

World Health Organization

E

Viale delle Terme di Caracalla, 00153 Rome, Italy - Tel: (+39) 06 57051 - E-mail: codex@fao.org - www.codexalimentarius.org

Agenda item 7 **CX/MAS 20/41/9 Add.1**

JOINT FAO/WHO FOOD STANDARDS PROGRAMME
CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING

REVISION OF THE GENERAL GUIDELINES ON SAMPLING (CXG 50 – 2004)

Comments at in reply to CL 2020/27-MAS

Comments of Australia, Canada, Chile, El Salvador, European Union, Iraq, Japan, Mauritius, Morocco, New Zealand, Norway, Peru, Thailand, United States of America, CCTA, EURACHEM, IUFOST

NOTE: CCMAS41 has been postponed to 17 – 21 May 2021. The comments compiled in this document will be made available to the EWG chaired by New Zealand and co-chaired by the United States of America for further consideration and preparation of a revised version of the Guidelines for consideration by CCMAS41.

Background

1. This document compiles comments received through the Codex Online Commenting System (OCS) in response to CL 2020/27-MAS issued in March 2020. Under the OCS, comments are compiled in the following order: general comments are listed first, followed by comments on specific sections.

Explanatory notes on the appendix

2. The comments submitted through the OCS are hereby attached as **Annex I** and are presented in table format.

ANNEX I

GENERAL COMMENTS	MEMBER/OBSERVER
Australia would like to thank New Zealand and the USA for their continued efforts and further development of the Proposed Draft Revised General Guidelines on Sampling (CXG 50) December 2019 document. Despite revisions to the EWG CXG 50 December 2019 version of this document, our comments as submitted at that time largely remain applicable and where relevant have been repeated below. General comments: 1) Two major new concepts have been introduced into this document including: a. Producer's Risk Quality (PRQ) and the Consumer's Risk Quality (CRQ) – which we can only find elsewhere in ISO 28592:2017 'Double sampling plans by attributes with minimal sample sizes indexed by PRQ and CRQ'. The objective of this International Standard is to provide procedures that enable lot disposition to be determined quickly and economically if quality is particularly good or bad. Followed for intermediate quality, a second sample drawn in order to be able to discriminate more reliably between acceptable and unacceptable lots. Is this ISO 28592:2017 specialised sampling plan the underlying objective for this guidance document? As we see suggestions of this throughout the guideline without a direct reference, and have concerns about this specialised sampling plan and its suitability in the broader international food trade context. Particularly when the sampling plans in ISO 2859-1 can be used for the inspection of lots in isolation, but the user is advised to consult the operating characteristic curves to find a plan that will yield the desired protection. b. Fractional Non-Conformance (FNC) plans – suggested where Variable sampling plans (classically ISO 3951) are applied to characteristics which do not follow a normal distribution. The term 'Fractional Non-Conformance' we could only find elsewhere, in the 'Codex Sampling 15 Feb 2019NZ' and papers Zhou et al (2018), Govindaraju et al (2015) . The 2018 paper states "The term fractional nonconformance refers to quantification of measurement uncertainty in probabilistic terms (as against providing a confidence interval for the true value using the repeatability standard deviation)….. In contrast to the traditional attribute method of classifying a measurement as conforming or nonconforming, the fractional nonconformance approach assigns a probability of nonconformance to specifications. The closer an apparent measurement to the specification limit, the higher the fractional nonconformance." Again, a specialised sampling plan for which we have concerns about suitability in the broader international food trade context without extensive user acceptance testing. Further explanation as to why these major new concepts were required, and how they improve the sampling process may be required. ********** 1 Xin Zhou, Kondaswamy Govindaraju & Geoff Jones (2018)" plus "Govindaraju, K., and Jones, G. 2015. Fractional	**Australia**

acceptance numbers for lot quality assurance. In Frontiers in Statistical Quality Control, vol. 11, pp. 271–286. Springer 2 Govindaraju, K., and Jones, G. 2015. Fractional acceptance numbers for lot quality assurance. In Frontiers in Statistical Quality Control, vol. 11, pp. 271–286. Springer	
In appendix IV - Control of Producer's Risk (PR) - 2nd paragraph The concept of producer's risk. In the zero-acceptance section of the outline, states that "When applicable, also check the level of non-compliant units that will be accepted with high probability under the plan and determine if this level is achievable by the supplier using GMPs or may result in too many rejections of acceptable lots." This indicates a check on producer's risk in some way, though it's not explained precisely how, and there's no mention if it as part of the zero-acceptance sampling plan itself. Without additional information on the parameters chosen for current food commodities, it is difficult to say how we should control for producer's risk. In appendix III - last bullet before table. There are limitations of a table that only looks at consumer's risk, such as the zero-acceptance table. The OC curve approach is more feasible when controlling for both PR and CR. On the other hand, the OC curve method proposed by might not be needed when only controlling for consumer's risk. The USA has provided a table for zero-acceptance at different sampling sizes which is easy to understand, though it only controls for consumer's risk for zero-acceptance plans. A general note related to the two different perspectives from the USA and New Zealand. New Zealand and USA do not necessarily contradict each other in terms of content, but their different approaches to the document clearly target different objectives as to its use. USA: PROS: Focuses on specific criteria used in current practices. More clearly lays out a set of instructions to steer sampling plans toward a specific set of procedures. CONS: May be too focused on most widely used sampling plans currently in food commodities. Document would provide less general information on sampling plans, and may not reflect sampling plans for projects undertaken in the near future under different conditions. New Zealand: PROS: Offers more information on a wider variety of sampling plans. Useful as a general reference and training document. CONS: Contains suggestions that are better for a textbook and may not be applied in practice. There may be too much	**Canada**

information in their draft that will not be used, either because they are not encountered for food commodities, or too complex to be used in practice. There may not be enough guidance on situations in which each procedure is most suitable/not suitable. Canada supports the idea of simplifying and focusing the document by including only the information that commodity committees, government and industry need to prepare sampling plans. Consideration should be given to the US outline as a starting point. Consideration of fairness could be achieved by including producer and consumer risk in the guideline. However, this would only work if the concepts and tools are explained in a simple and clear manner.	
Regarding agenda item 7, Chile wishes to express their appreciation to New Zealand and the United States for the work carried out in the Electronic Group. We consider that several points of the document have been duly clarified, however, there are a series of references that emphasize ISO standards, which in our opinion is correct, but the document should be more explicit as regards the indications so that it can be a more useful tool for laboratories and inspection bodies.	**Chile**
El Salvador appreciates the document submitted by the Codex Alimentarius Secretariat prepared by the Electronic Working Group led by New Zealand and the United States of America. El Salvador has analyzed the matter in question in Step 3 in the National Codex Mirror Committee on Methods of Analysis and Sampling and presents the following comments to Annex I "Proposed Draft Revision of the General Guidelines on Sampling (CXG 50-2004)" and Annex II "Preliminary draft of an electronic book" to answer in circular letter CL 2020/27/OSC-MAS. General comments: • El Salvador agrees with the Preliminary Draft and believes that the structure and content of the guideline provide a simplified and understandable guidance on the design and selection of sampling plans that best suit the user. It is seen that other Codex standards related to the guidelines are validated and mentioned, as well as other international standards to avoid the duplication of information. We agree with the progress of the Preliminary Draft presented in Appendix I in the process of elaboration of Codex Alimentarius Standards.	**El Salvador**
The European Union and its Member States (EUMS) congratulate the eWG under the lead of New Zealand and the co-lead of the United States of America for the further development of the revision of CXG50-2004. Sampling is an essential element for the verification of provisions in Codex standards and the current version of CXG50 is a useful information source, though sometimes difficult to read and understand for non-specialists. Therefore, a revision o CXG50 with the overall aim to simplify its structure and language to provide effective guidance to all CAC subsidiary bodies and interested parties for designing/selecting sampling plans is highly welcome. The EUMS are of the opinion that the presented draft succeeded in principle in reaching the goals set out by CCMAS39 and CCMAS40 in providing not only guidance in the main document but also technical solutions in form of an e-book. The presented draft, however, could benefit from: • Avoiding redundancy, e.g. the aspect of ‘fair and valid sampling procedures’, though of highest relevance, is	**European Union**

over emphasised throughout the text; • clarifying the relationship between relevant ISO standards (i.e. ISO 2859 and ISO ISO 3951) and CXG50, in particular as it relates to the presented roadmap (see 4.1.1); • focusing on the most common scenarios encountered when developing/selecting sampling plans and providing more detail to help users of the guidance to select the correct approach related to attribute plans for sampling from discrete items/units and variable plans for sampling of bulk materials. In particular, more information regarding composite sampling (e.g. number and amount of increments, homogeneity of composite sample, sub-sampling, test portion size, etc) will be helpful. Background and plans for niche applications could be explained in the e-book. The EUMS are aware that the proposed CGX50 text follows a different approach and style compared to the current version, which was the intent of the re-draft. Information provided as tables and figures in the current CGX50 will be moved to an e-book and will be provided in form of software apps, which is an appropriate and smart way of helping users to understand the impact of certain plans. NZL and USA as chair and co-chair are invited to inform CCMAS41 whether the software has been validated and how maintenance and access to the software can be ensured in a sustainable manner.	
we agree with proposed draft revision of the general guidelines on sampling (CXG 50-2004).	**Iraq**
Japan would like to thank New Zealand and the United States for all of effort in preparing the discussion documents about revising CXG 50-2004. To our knowledge, the aim of revising GXG 50-2004 is to improve user friendliness, not to review the full content of existing CXG 50. The efforts for revising CXG 50 should be made to provide a simpler more understandable guideline in line with the new work proposal(REP18/MAS-Appendix V). If a full review of CXG 50 from the viewpoint of statistical sampling plan was proposed as new work, it might not be sure whether the proposal was approved in CCMAS. In the EWG for CCMAS40, Japan suggested that the structure of draft revised CXG 50 should be the same as that of the existing CXG 50 as much as possible in order to be user friendly. However, the proposed draft CXG 50 is an entirely different document that misses much of the practical guidance and useful information as mentioned by the US. Japan suggests that further discussion will be held with due consideration of US proposal and that the US proposal "US proposed CXG50 revision Top-level Outline" shown in Appendix III to CX/MAS 20/41/9 will be appropriate structure of revised CXG 50.	**Japan**
Mauritius thanks New Zealand for all its work and effort in the preparing the circulated document and the US for their comments. We also agree that the subject area under consideration in the circulated draft is complicated and requires prior knowledge of sampling and statistics and the ISO documents on statistical sampling. Taking into consideration the fact that CXG50 needs to be updated, we fully support continuous improvement of the existing draft. From our part, it is being suggested that more practical examples could be provided (with figures from real data) in the current draft to make it more accessible to all the potential users of this document. For example, at page 5, paragraph before last: we could provide more examples there. We support US's comments that "practical guidance and useful information" be retained; that end-product testing is not	**Mauritius**

an end in itself but should be used in conjunction with appropriate food safety management systems. We also concur that the preamble “should be understandable to audience with limited statistical training and relatively jargon free” On the other hand, The concept of measurement uncertainty is relevant in the context of sampling. ICMSF is very useful and applicable in the field of sampling in food microbiology. We do not agree that guidelines should mainly focus on attributes plans.	
Morocco is in favour of developing sampling plans by specific areas dedicated to products or groups of products sharing more or less common characteristics in their presentation, the objective of the revision being to make the general guidelines on sampling (CXG 50-2004) simple to apply and easy to understand with clear examples.	**Morocco**
In accordance with the direction provided in the CL, New Zealand (NZ) submits comments on the proposed draft revision of the General Guidelines on Sampling (CXG 50-2004) and the e-book, uploaded to the Codex Online Commenting System (OCS): https://ocs.codexalimentarius.org/. In submitting comments, NZ considered the information provided in CX/MAS 20/41/9, and the conclusions and recommendations in paragraphs 21 – 24. Our understanding is that the OCS process provides for comments on specific paragraphs of CXG 50 and general comments or summary comments on the document as a whole. Contents Request for comments on the proposed draft revision of the General Guidelines on Sampling (CXG 50-2004) 1 Background 1 General comments on the revised CXG 50 3 Strategic direction revised CXG 50 3 Motivation for the revised CXG 50 3 General comments on the Supplementary e-book 4 Expanded points for consideration from Appendix IV of CX/MAS 20/41/9 4 Attribute Plans 4 Focus on attributes plans 4 Zero Acceptance Number (ZAN) attributes plan 4 FAO/WHO and ICMSF Plans 5 Inefficiency of attributes plans 5 Distribution free plans 5 Measurement error 6 Reinspection 7 Additional points for consideration from Appendix III of CX/MAS 20/41/9 7 ISO Standards 7 ISO terminology 7 ISO plans (schemes) 7 The level of understanding needed for the revised CXG 50 8	**New Zealand**

Specification of risks as inputs to plan design 8 Including only attribute plans 8 Operation of variables plans 9 Homogeneity 9 'Risk' terms 9 Acceptable consumer risk 9 Specifying only a limit does not define acceptable risk 10 General comments on the revised CXG 50 Strategic direction revised CXG 50 NZ supports the ongoing progress for the revised CXG 50 based on the process to date involving the agreed project document , the prioritisation list , the EWG consultation and the development of the supporting e-book containing the apps. NZ acknowledges the USA comments but proposes a compromise between the revised CXG 50 and the USA proposals to progress the revision. This compromise involves leaving the structure of the revised CXG 50 as is, with its current level of focus on attributes plans recognising their fundamental role in sampling but also including additional topics on the prioritisation list as well as those areas that have been proposed by the EWG (such as the sections on physical sampling and random sampling). Some prioritisation list topics resulted in the inclusion of a wider range of sampling options. The proposed compromise would allow users, particularly national regulatory officials, to design valid sampling plans appropriate to their situation, using whatever sampling plan options that are available and taking measurement error into account where applicable as well as economy of sampling and testing. However rather than specifying sampling plans themselves, Commodity Committees would specify outcomes in terms of maximum allowable consumer's and producer's risks for sampling plans, in the same way as those risks are specified in International Recommendation OIML R87 for the inspection of weights using the Average Quantity System. This proposed compromise would be consistent with the Purpose of Codex Methods of Sampling 'to ensure that fair and valid sampling procedures are used when food is being tested for compliance with a particular Codex commodity standard' as it provides for Commodity Committees to specify outcomes for sampling plans for Codex provisions. Motivation for the revised CXG 50 We note comment about the relevance of sampling 'how routine sampling should only be used when no better tool is available to assure objectives are met'. Options such as process control are available only to producers [can we say why the producers would sometimes use routine sampling?]. Consumers have the prerogative to check compliance of the product which they can only do using sampling. However there is no standardisation of sampling requirements, or how test results are to be interpreted, so practice varies. Audits and similar activities might be carried out, but these are essentially subjective. As outlined elsewhere, process control is not always used or feasible and producers must often rely on sampling. The current lack of standardisation in sampling inspection by regulatory agencies and consumers is a key motivation for the development of Codex guidelines, to standardise implementation of inspection of provisions in Codex	

standards. In addition, the revision will provide for a 'user-friendly' approach that provides a simpler way to understand and develop statistical sampling plans. It does not typically refer to a 'simple sampling plan' approach. General comments on the Supplementary e-book We also note some comment on a preference to use a simple Excel spreadsheet to replace many of the existing tables in CX G50. The e-book has been developed to provide a practical and usable resource to support the revised CX G50. The e-book includes information, examples and apps. The use of apps is a more modern approach and enables less statistically educated user's access to other, more complex types of plan for which calculations cannot be easily performed in Excel. The suggested use of an Excel spreadsheet presumes that the options for sampling plans presented in the revised CXG 50 will be limited to the simplest cases allowing them to be implemented in Excel. Expanded points for consideration from Appendix IV of CX/MAS 20/41/9 Attribute Plans Focus on attributes plans NZ believes that widespread use of certificates of analysis for trade and general lack of awareness about other sampling methodology and measurement error has brought about and encouraged the classification of results as pass or fail with respect to limits and the use of attributes plans. We believe there are other valid options which should be included in the revised CXG 50, along with guidance on their use – where applicable, variables plans offer a way of providing more stringent assessments with more economical sample numbers. We note that there are already sampling plans in Codex, such as those for aflatoxins, that are not based on inspection by attributes. In NZ's view, basing the revised CXG 50 only on attribute plans and requiring that only plans from CXG 50 should be used is too restrictive, to the point of being prescriptive, and is likely to impose unnecessary costs on consumers and producers in respect of sampling and testing. Zero Acceptance Number (ZAN) attributes plan NZ believes that use of these plans is driven by the perception that if non-zero acceptance numbers (c>0) are used, lots containing non-conforming product will be allowed to pass, consistent with the usual interpretation of test results on certificates of analysis. NZ believes that the widespread use of these plans is driven by their use in food safety and metrological applications in which non-conforming items in the sample are not permitted in accepted lots and by the perception that if non-zero acceptance numbers (c>0) are used, lots containing non-conforming product will be allowed to pass, which is consistent with the usual interpretation of test results on certificates of analysis. However this perception is incorrect, as no matter what sampling plans are used, there will always be a chance of accepting lots containing unacceptable levels of poor quality product. The design of sampling plans needs to be done to control the risks to suitable levels. Use of ZAN plans may be unsuitable for Codex commodity defects and other provisions as they do not explicitly control producer's risks. We note further that these plans provide poor discrimination between good and poor quality lots. FAO/WHO and ICMSF Plans Sampling plans presented in FAO/WHO and ICMSF and possibly those for pesticides in CAC/GL 33 are special cases	

and may not be suitable for general use. These plans rely on the use of tightened limits for control of consumer's risks and economy of sampling so may be unsuitable on the grounds of potential unfairness to producers.
The general approach to design of plans for non-food safety applications should explicitly (i.e. specified in the design) control both producer's and consumer's risks to ensure plans are fair.
Inefficiency of attributes plans
While attribute plans are fundamental in sampling, they can, as shown by the table for inspection of pesticide residues, be intrinsically inefficient if high levels of assurance are to be provided. In particular classifying measurements into binary [pass or fail] outcomes causes a loss of information on the variation within the lot, which can be recouped only by taking large numbers of samples. However use of large sample numbers is not the only means of providing high levels of assurance – measures such as tightened limits can be used although this may be unfair provides considerable motivation to consider other, more economical options for sampling plans. These options are described in the revised CXG 50.
While attribute plans are fundamental in sampling, they can, as shown by the table for inspection of pesticide residues, be intrinsically inefficient if high levels of assurance are to be provided. In particular, classifying measurements into binary [pass or fail] outcomes causes a loss of information on the variation within the lot, which can be recouped only by taking large numbers of samples. This table provides considerable motivation to consider other, more economical options for sampling plans. These options are described in the revised CXG 50.
For example we include details of sampling plans for inspection of a lot in which the Producer's Risk (PR) is assumed to be 5% (5% rejection) at the Producer's Risk Level (PRQ) of 5% non-conforming and a Consumer's Risk (CR) of 10% (10% acceptance) at a Consumer's Risk Level (CRQ) of 10%. We note that not all of these options will be available in all situations.

Plan Type Tested	Samples	Acceptance constant
Attributes	52	c=2
Variables (unknown standard deviation)	20	k=1.74
Variables (known standard deviation)	8	k=1.74
Compositional Fractions	1	TBD*

* Details of the acceptance criterion of the sampling plan for mass fraction parameters, specifically the number of increments needed to form the composite sample, depend upon the precision parameter for the quality characteristic concerned (as well as the specification of allowable risks).
Distribution free plans
The use of attributes plans for situations where the underlying distribution cannot be satisfactorily described by a suitable distribution has already been included in the revised CXG 50. However we feel that general restriction to attribute plans is unnecessary as there are other valid options available in many cases.
Consignments, being collections of lots or part lots, are not necessarily inhomogeneous, it will depend upon the source of the raw materials, the final product, processing and other issues.
It should be noted that the distribution employed serves only as a model for the behaviour of the characteristic in a lot, it need not be 'exact' – we are not fitting a distribution to the data as such, we are using the assumed distribution to make a decision on lot acceptability.

Measurement error With reference to the comments about attributes plans, we believe that there are several reasons measurement error might not be allowed for: - Lack of awareness about measurement error, it being common to classify results as pass or fail with respect to limits. It is very common to consider test results as what is in the product but this does not allow for the effect of measurement error. - Lack of knowledge of how to make allowance for measurement error, which requires some statistical knowledge. In particular, it does not seem possible to allow for measurement error in attributes plans. Measurement error has the potential to substantially alter the control of risks in sampling [refer App15] and cannot always be ignored; it might be invalid and unfair not to do so. We also note: 1. There is no standard Codex procedure for how limits are derived and it is not clear how results have been, or are, used to set limits 2. It is logical that even though results may have been examined in the limit setting process, those limits apply to the product. Indeed this is the standard way the statistical literature (including ISO3951-6) treats limits, as relating to the product 3. Ignoring measurement error would conflict directly with CAC / GL 54 - 2004; the objective of conformity assessment is to determine whether the true value of the sample complies to the limit [for the product] 4. We are aware on one situation, for the Protein-Solids-Non-Fat ratio in milk powders, where the limit was NOT set as the minimum or the maximum of the data, so the measurement of all laboratories is not necessarily taken into account in the limit setting 5. In any case the data examined during the limit setting process is probably not sufficient to reflect the extremes of measurement error that, unless negligible, may exist across all labs and future inspections. Measurement error is considered significant if the 'error-variance ratio', the square of the ratio of the measurement error standard deviation and the standard deviation representing the variation in the product exceeds 10%. Normally allowance for measurement error would be made in the sampling plan (i.e. the acceptance criterion and the number of samples required) if this ratio exceeds 10%. If we already had information on the ratio it could be used directly in the appropriate app (App15) to design a plan to assess whether allowance should be made for measurement error in the sampling plan. NZ believes that the significance of measurement error in sampling needs to be evaluated on a case by case basis, taking account of the variation of the characteristic of interest in the product and the measurement error of the relevant analytical method. Reinspection Our consultant statistician, an expert on sampling, recommended inclusion of this section as the guidelines lacked advice on measures to be taken when disputes over sampling occur. This procedure is used only when the outcome of inspection is considered suspect for some reason (do we need a for example?); with a specified number of reinspections allowed. We appreciate that the immediate reaction of many readers is that reinspection is not valid as it provides greater opportunity for lots to pass. However in practice it provides better control of producer's risk and of failing lots of good quality as poor quality lots will likely fail reinspection. Thus, the use of reinspection is justified in the same way as sampling schemes, sets of sampling plans with switching	

rules, are preferred over the use of isolated sampling plans; to provide greater surety that the correct decision, of accepting or rejecting a lot, has been made as inspections made using relatively low sample numbers carry high risks of making incorrect decisions App6 in the e-book would be used to evaluate resampling [reinspection] options in the event that an assessment of a lot was challenged and it was agreed that resampling of that lot would be undertaken. Alternatively App6 could be used to evaluate resampling options as part of the agreement between parties prior to the commencement of trade. Additional points for consideration from Appendix III of CX/MAS 20/41/9 Further to some of the other points raised by the USA we provide some explanations below to help with understanding. ISO Standards ISO terminology This is the standard terminology which we are reluctant to remove, doing so will cause the revised CXG 50 be inconsistent with, and to become removed from international standards and other material on the subject. We believe the proposed process, of specifying BOTH risks in the design of sampling plans will be more transparent and provide users of a greater understanding of how the design process operates and what those parameters mean. We note that ISO has recently changed the terminology to Producer's Risk [Point] and Consumer's Risk [Point] which are expected to be more meaningful to users. ISO plans (schemes) Several comments have been made around the inappropriateness of the ISO plans. NZ agrees - the material on the ISO plans has been included in the revised CXG 50 only to explain why that approach has not been followed in CXG 50, so that both consumer's risks and producer's risks to be controlled explicitly in situations where both risks should be considered such as applications other than those relating to food safety, the inspection of commodity defects and macro composition. Adoption of this approach provides greater transparency in the design of sampling plans so that there should no longer be confusion arising from the use of the wrong type of plan. We also note that the distinction between plans for isolated lots and continuing series of lots is an outcome of the way ISO plans have been designed - but this distinction is no longer relevant if both consumer's and producer's risks are considered in the design of sampling plans. The level of understanding needed for the revised CXG 50 One key role of the revised CXG 50 is to provide accessible guidance on other types of plans to users, however those users will still have to invest some time to become familiar with the new information. The use of apps means that users need only understand the concepts. 'User-friendly' in this regard refers to an approach that provides for a simpler way to develop and understand statistical sampling plans. It does not typically refer to a 'simple sampling plan' approach. NZ believes that the inclusion of apps for various plans and for the demonstration of principles, makes this process much simpler – design and evaluation of sampling plans can be undertaken with an understanding of the basic principles without having a background in statistical theory. Specification of risks as inputs to plan design The levels such as '95% confidence for accepting good lots, 90% confidence for rejecting bad lots' are used only in the design of a sampling plan which requires the specification of two points on the intended Operating Characteristic	

curve. However it is important to examine how the resulting sampling plan controls the consumer's and producer's risks over the entire range of levels non-conforming - in practice the consumer's and producer's risks vary according to the level non-conforming in a lot. The choice of the two points for the design of a plan is essentially arbitrary although it is usual to specify the PRQ as the quality level at which there is a Producer's Risk of 5% rejection and the CRQ as the level at which there is a Consumer's Risk of 10% acceptance. These risks, the probabilities of rejection and acceptance can be set arbitrarily but this complicates the design process as one is then faced with varying four parameters rather than just two. This seems unnecessary as for the same plan a 10% chance of acceptance at the CRP corresponds to a 5% chance of acceptance at a different level (greater than the CRP). We have updated the revised CXG 50; in the explanatory text beneath the OC curve diagram in Section 2.2.1. 'Confidence' levels of 95% and 90% are not recommended, the choice of these levels is arbitrary, but use of these is conventional, so that users need be concerned only with setting the PRQ and CRQ quality levels. Including only attribute plans NZ believes that restricting the revised CXG 50 to include only attribute plans is unnecessary as there are other valid options available in many cases. Consignments are not necessarily inhomogeneous, it will depend upon issues such as the nature and source of the raw materials, the amount of processing and the level of process control It should be noted that the distribution employed serves only as a model for the behaviour of the characteristic in a lot, it need not be 'exact' – we are not fitting a distribution to the data as such, we are using the assumed distribution to make a decision on lot acceptability. Fitting a distribution to the data can only be done so well with the limited data that is usually available. However an assumption of an underlying distribution is often needed for the evaluation of plans, such as the log-normal distribution for evaluation of 3-class attribute plans. Operation of variables plans The comments suggest an assumption (as in FAO/WHO) that standard deviations used in variables plans are considered known, i.e. that their true values are known, such as the assumption that σ (sigma) = 0.8 for Aerobic Plate Counts in ground (minced beef), used in the evaluation of some micro 2- and 3-class attributes plans. However this is not what the revised CXG 50 is advocating: - Known standard deviations are used only when there is evidence that the underlying 'process' generating the results, is in-control; standard deviations would be based on long term performance of that process - Consumers would not normally be in a position to know the true values of standard deviations to use in assessments - Variables plans based on unknown standard deviations, the so-called "s-method" can be used estimating standard deviations from the inspection data itself, but at a cost of having to take more samples - The use of unknown standard deviations is self-compensating, increases in the observed standard deviations require corresponding reductions in the observed average levels for lots to be accepted; increases in the observed standard deviations do not mean increased levels non-conforming in accepted lots Homogeneity This section is still under development in the revision of the guidelines. While it is not necessary to assume the characteristic in a lot follows the same distribution, the statement is not entirely true - App5 demonstrates the effect of	

lot heterogeneity in the form of correlation - on the control of risks when attribute plans are used. Homogeneity means that the product is of the same quality throughout the lot, i.e. with the same level non-conforming throughout. Assumption of normal distribution implies homogeneity; homogeneity is about the characteristic having the same distribution throughout or more specifically the 'same' level non-conforming. 'Risk' terms We refer ISO9000:2015 from which it is apparent that food safety risk is just one aspect of risk in general. There are different interpretations of 'risk'. Codex needs to take the wider view, rather than immediately associating the term risk with food-safety – the term risk is the accepted terminology used in risk management standards. Our proposed action is to clarify inconsistency of wording describing risks in the next version of the revised CXG 50. Acceptable consumer risk Sampling plans do define acceptable risk – this is precisely how, as we understand it, they are used in microbiological criteria normally defined as a sampling plan as an outcome to be achieved i.e. if a consumer tested the product using that plan. However this is widely misunderstood with producers believing that they should use this same plan (against our advice that it is often not appropriate for producers to use the same plans as those used by consumers). FAO/WHO provides a definition of microbiological criteria: 'A microbiological criterion is a risk management metric which indicates the acceptability of a food, or the performance of either a process or a food safety control system following the outcome of sampling and testing for microorganisms, their toxins/metabolites or markers associated with pathogenicity or other traits at a specified point of the food chain'. Specifying only a limit does not define acceptable risk The revised CXG 50 approach is to use the accepted standard statistical approach for sampling plans, defined by specifications of acceptable risk both from the consumer's and the producer's perspectives. It is for the commodity committees to decide upon acceptable levels of risk, the purpose of CXG 50 is to provide tools to help them decide. The revised CXG 50 should not include recommendations on what constitutes acceptable risks, nor should it be written in a way that attempts to give guidance to or pre-empts decisions made by commodity committees.	
Norway would like to thank New Zealand for their extensive and impressive work in drafting the document, reviewing, and responding to all received comments. With reference to CX/MAS 20/41/9 and the recommendations in paragraphs 23 and 24 we would like to submit the following general comments: Based on comments from the co-chair of the EWG, Norway finds it important to prioritize finding consensus on the overall structure rather than forwarding specific comments on the new draft. This would in our view, be more in line with the purpose and the aim of the revision; How to make the CXG50 more user-friendly for the target group; bearing in mind that they do not necessarily possess knowledge and interest in statistics, and may not be familiar with the different ISO standards for sampling plans. We would like to support further development and validation of apps using links provided in the e-book.	**Norway**

We support focusing on attribute sampling plans, and hence, measurement uncertainty would not be a topic in this document. However, we find measurement uncertainty important in other contexts since it is a property of every measurement made and is needed when comparing measurements and when checking for compliance with a specification; if the analytical results (with or without the measurement uncertainty) is within the specified limits. We support the proposed CXG50 Revision Top-level Outline as it is processed based. The outline follows the process of what is to be considered in planning the sampling for specific purposes, rather than explaining different sampling plans and leaving the reader to choose which plan to use. Much of the content covering the different topics is available in the current CXG50 and in the draft prepared by New Zealand. We would welcome continuation of the work as outlined by the USA.	
1. Comments on the Draft Revision of the General Guidelines on Sampling (CXG 50-2004) and the electronic book (Prepared by the EWG led by New Zealand) - Comments at Step 3. Peru thanks the electronic working group led by New Zealand for the effort undertaken in the revision of CAC/GL 50-2004 and the drawing up of the distributed draft, which gives us the opportunity to submit the following comments. Peru in general agrees with the technical concepts expressed in the Preliminary Draft, and we are in favor of including a support guide as proposed by the Electronic Working Group. However, we think that the distribution/structure of the Preliminary Draft should be improved in order to help commodity committees as well as regulatory entities in Codex member countries and other users to understand the principles of sampling. In this sense the document should provide practical examples that could be added in each chapter of the document or in an annex to it. 2. Comments on the points raised in paragraph 24 of document CX/MAS 20/41/9. a) The key technical areas identified by New Zealand in response to comment from the United States (Appendix IV): • Regarding the approach focusing on the attribute plans: We believe that the CAC/GL Guidelines should provide information on the various sampling plans. It could be considered to give more emphasis to the attribute sampling plans due to their high level of applicability, but without stopping to explain the variables plans. • Regarding the relevance of the 2016 FAO/WHO Guide: We think that the FAO/WHO document "Statistical aspects of microbiological criteria related to foods: a risk managers	**Peru**

guide" should not be used as a model for the design of sampling plans. Sampling plans should be based on the standard statistical approach presented by Codex Guidelines CXG 83-2013 Principles for the Use of Sampling and Testing in International Food Trade (GL 83) in which sampling plans are designed based on the specification of producer's and consumer's risks. We are in favor of taking advantage of technological progress to build correct sampling plans. • Regarding the margin for measurement error: We are in favor that the CAC/GL 50 Guidelines should provide the necessary guidance regarding measurement error and its adjustment. It is suggested that examples on this point be provided in the Guidelines. • Regarding control of producer's risk: We agree that the CAC/GL 50 Guidelines should maintain the standard statistical approach and consider producer and consumer risk in the design of sampling plans. Furthermore, we agree that the objective of the CAC/GL 50 Guidelines is to provide a set of tools to assist commodity committees and others with designing appropriate plans for each provision in Codex standards. In this regard, decisions on acceptable risk levels are matters for Codex commodity committees to decide. b) The key technical areas identified by the United States (Appendix III): • Concerning that most of the sampling plans used in Codex are based on attributes sampling: We think that the draft CAC/GL 50 guidelines should not only consider attribute sampling but also present other sampling options, such as variable sampling, etc. • Regarding that it is not necessary to take into account the measurement uncertainty: We consider that, depending on the case, measurement uncertainty is be taken into account to make decisions for the declaration of conformity. • Regarding that the material on re-inspection should not be included in the Guidelines: We believe that re-inspection should be included in the Guidelines, and should be taken into account that re-inspection is applicable many times by health authorities and conformity assessment entities in scenarios such as fishery and agroindustrial inspections. • Regarding other points indicated by the USA: We agree with the comment of US that routine sampling should be performed when the supplier does not have a tool	

that allows achieving food safety and quality objectives through the HACCP system or other production control systems. c) Concerning the proposal of the United States for a general design of the revised Guidelines (Appendix III): In general, we are in favor of the general design proposed by the United States regarding the revision of the CAC/GL 50 Guidelines, except of what is mentioned in Chapter 8 References, that needs more consideration. In our view, references in the Revised Guidelines should not be limited to citing only FAO/WHO publications. ISO and ICMSF publications are also considered sources with a standard statistical approach. It is suggested to include in the references the most important ISO and ICMSF publications that agree with the FAO/WHO publications.	
1. Thailand support to continue the work on the updating/revising of CXG 50. However, in our opinion, the current CXG 50 – 2004 provides clear technical information and comprehensive contents with examples that are useful for users to consider and understand sampling plans. Therefore, the current CXG 50 - 2004 should be retained and used as a basis for the revision. The revision of CXG 50 should be conducted by providing additional information/contents or revising the document for relevant areas. 2. The revised CXG 50 should provide ready - to - use tables of examples of sampling plans for Codex Commodity Committees to easily choose the appropriate sampling plans for each commodity. The mentioned tables should be brought from Information Document: Practical examples of sampling plans that provides examples for choosing a sampling plan for each situation. CCMAS should kindly convert the sampling plan tool from e-book into ready - to - use tables of sampling plans that vary by different key parameters, including producer's and consumer's risk that will be considered by CCMAS in the consideration of approving sampling plans. This will assist the Commodity Committees to choose and apply the given table of sampling plans by choosing parameters that are suitable and fit for their products being considered. Other than facilitating the Commodity Committees, this will decrease a problem that the Commodity Committees do not consider all parameters required by CCMAS. 3. The revised CXG 50 should identify a clear scope that it covers acceptance sampling plans and is applicable for only the control at reception, but may not be applicable for control of end - products and for process control during production.	**Thailand**
1. Regarding *CX/MAS 20/41/9, Appendix IV, "The New Zealand response to four key areas identified by USA"*: The four points identified by New Zealand were discussed, however broader comments are listed by importance below. 2. The work was expected to revise CXG50 for ease of understanding and use. However, rather than a revision, the proposed draft is an entirely different document that is missing most of the practical guidance found in CXG50. The draft repeats general sampling principles found in other standards, while guidance on plan design is unclear. Annex 1 is a model revised CXG50 Outline which was presented as Appendix III of CX/MAS 20/41/9 and is captured here so it can be considered in unison with Annex 2. Annex 2 is model text on "Basic Concepts of Sampling', for consideration by CCMAS during revision of CXG50. 3. The draft e-book contains extensive guidance that is the actual substance of the proposed draft. However, the	**USA**

e-book is too complex and impractical for most users. Most of the apps are unnecessary or controversial. Two or three apps covering attribute and variables-type plans would be appropriate. An Excel spreadsheet (as first proposed) is all that is necessary and would be more useful for Codex users. Information on plan type, purpose, design, and inputs and outputs (used in tools) should be in the body of the guidelines, with minimal repetition in the tool itself. 4. The proposed revision contains a supplier-oriented bias (also seen in ISO and other industry-oriented sampling literature) that is confusing and misleading to commodity committees and customers designing sampling plans for receiving-oriented situations. The guidelines should correct bias and confusion perpetuated from the literature, such as using 95% confidence for accepting good lots, and 90% confidence for rejecting bad lots. 5. The 'Purpose of the Guidelines' should clarify that sampling plans do not define 'acceptable quality' or 'allowable risk'. It should also make clear that food safety and suitability objectives are achieved through GMP and HACCP systems that have performance criteria that are more stringent than can be verified using practical sampling plans. 6. The guidelines should clarify that the definition of an unacceptable lot does not change during design of the sampling plan. Performance criteria that define an unacceptable lot are based on risk assessment and food safety/quality objectives. These criteria are determined before selecting a sampling plan that attempts to detect the unacceptable lots. 7. The guidelines should discuss how acceptance sampling should only be used when more effective tools are unavailable (e.g. HACCP control records). (See ICMSF 2002, *Microorganisms in Foods 7, Microbiological Testing in Food Safety Management*, Section 5.5.1.) 8. The guidelines should discuss acceptable risk levels for food safety and suitability hazards (e.g. one incident per thousand servings, one incident per million servings), and how these levels cannot be detected using practical sampling plans. 9. The guidelines should clearly show (using a table) the practical limitations of sample size. For example, a marginally practical, 29-sample, zero-acceptance number attribute sampling plan can only reliably detect (with 95% confidence) lots with at least 100 non-compliant units per 1,000 units. 10. The guidelines should contain a table for zero-acceptance number attribute sampling plans (see Annex 2). These commonly used plans are universally-valid, cover most user needs, and can often be appropriately selected from a table of options without the need of a sampling tool app. 11. The guidelines should replace confusing ISO jargon (e.g. AQL/LQ and PRQ/CRQ) with plain language. The Codex Committee on Fish and Fishery Products believed that the "Acceptable Quality Level" (i.e. AQL = 6.5) provided a degree of consumer protection, however it only protects the supplier. Some CCCF sampling plans also reference AQL = 6.5. 12. The proposed draft and e-tool guidelines on re-inspecting rejected lots is statistically invalid. Instead, the guidelines should warn against this malpractice (see FAO/WHO 2016 FAO/WHO 2016. *Statistical Aspects of Microbiological Criteria Related to Foods*, Section 2.2, page 31). 13. The Guidelines should discuss why 'attribute sampling plans' are widely used for trade lot inspection because they are 'distribution-free' and not affected by unknown distributions of the characteristic of interest. Trade lots often contain product produced on different lines, on different days, or from different raw material sources, which affects distribution both within and between lots. 14. The guidelines should explain that 'variables sampling plans' require a fixed stable standard deviation to be	

useful and reliable (i.e. continuous lot inspection of a single product within a production facility, or similar circumstance). For typical trade lot acceptance testing, it is inappropriate to assume that the standard deviation will remain constant from lot to lot.

15. Bulk material sampling does not require a separate section in the guidelines. The nature of containment (bulk or packages) does not affect lot homogeneity or the statistical design of sampling plans. It only affects how the random samples are physically drawn.
    The notion that bulk material requires a separate section may have arisen from ISO-10725 (*Acceptance sampling plans and procedures for the inspection of bulk materials*). This standard is specific for a continuous series of lots with a uniform standard deviation, and a lot acceptance provision based on a lot average. This special situation is equally applicable to packaged material and can be covered in general sections. Note, this situation is not applicable to trade lots with changing standard deviations, or when the safety or suitability of individual sample units (servings) is important.
16. The issues experienced by committees using CXG50 are primarily related to the references to ISO standards. These standards have several situational assumptions and subjective aspects that are not applicable to food lot acceptance testing. In addition, ISO standards were designed for industrial inspection of machine parts and assume users can test large numbers of samples. References to ISO standards are problematic and should be removed.
17. The ICMSF example cases (based on health risk) are incomplete and should be removed from the guidelines. They are based on presumptions, detailed in ICMSF publications, that are inconsistent with trade lot acceptance testing, e.g.:
    a. Presume continuous lot-by-lot inspection from same supplier in such a manner that even a very low probability of detecting unacceptable lots will cause the supplier to improve practices.
    b. Justify small sample sizes based on practical limitations of traditional facilities used for bacterial incubation and enumeration.
    c. Plan performance for the 3-class attribute plans is based on a 10-fold difference between acceptable and marginally acceptable limit values (“m” and “M”).
18. The guidelines should avoid the term ‘risk’ for describing the statistical probability associated with sampling. Readers of Codex standards associate ‘risk’ with ‘health risk’. The proposed draft uses the term ‘risk’ loosely and in a way that readers will associate it with human health risk instead of mathematical probability.
    There are four traditional terms used in the sampling literature, “consumer’s risk”, “producer’s risk”, “consumer’s risk point” and “producers risk point” that are unavoidable. Therefore, these four (non-health) ‘risk’ terms should be allowed in the guidelines and be clearly defined and kept in quotes to prevent confusion.
19. Codex commodity standards and sampling guidelines are predominantly used for border inspection and other receiving oriented situations. The guidelines should not attempt to cover ‘statistical process control’ (an advanced subject), or ‘end-product testing’ (an ineffective control).
20. Measurement error is rarely incorporated into sampling plans, and it is inappropriate to incorporate measurement error into Codex sampling plans. Limits in Codex standards make allowance for the measurement error of the reference method, or method criteria. Measurement error should not be accounted for twice, both in the limit and in the sampling plan.
    Note that measurement error is usually insignificant relative to the variation between sample units, and that safety factors in health-related limits are orders of magnitude greater than measurement error.

21. The guidelines should discuss methods of random sampling, and discourage improper practices, such as combining lots prior to sampling. 22. Composite or aggregate samples are used for different purposes (unrelated to bulk or packaged material) and these should be categorized and explained in the guidelines. 23. Virtually all hazards and defects are heterogeneously distributed in food lots. CXG50 incorrectly indicates that lot homogeneity is required to use sampling plans. The revision should clarify that lot homogeneity does not affect the applicability or performance of attribute sampling plans, which are distribution-free. Lot homogeneity also does not affect variables sampling plans for lots that fit the presumed distribution (e.g. normal) and dispersion parameter(s) (e.g. standard deviation) that accounts for the degree of heterogeneity. 24. In accordance with CXG50, the scope of the revision should not cover double, multiple, sequential, and other complicated sampling schemes, which are often misapplied. 25. The guidelines should discuss different sampling objectives and how these relate to sampling frequency and stringency. 26. The guidelines should discuss the relevant sample unit amount, and the advantages of many smaller sample units over fewer larger sample units (see FAO/WHO 2016, Example 14, p. 48). 27. The requirements for CCMAS approval of sampling plans should be listed in the Codex *Procedural Manual* and not elaborated differently in CXG50. 28. Microbiological sampling is adequately covered in CCFH standards and FAO/WHO 2016 (which should be referenced). 29. The guidelines should capture both practical and theoretical aspects. For the average readers/users the inclusion of extensive formulas acts as a detriment to understanding and following the text. However, for technically trained readers/users those formula offer explanation and insight into the guidelines and specifically how the APPS work. While not in the main body of the guideline, but perhaps in an Appendix or Annex, the formulas used in the APPS and which support the text should be presented. 30. For some apps, technical information (formulas) in varying levels of details are provided in supplemental materials or notes. In some cases, the apps appear to work like “black box”, though most of the underlying statistical theories (formulas) are simple, often just the Type I and II errors calculated as the tail areas of distributions. The levels of technical detail need to be made consistent across apps, and the information should be self-contained. This technical exposition should include, in addition to the statistical theory behind each app, (1) the algorithm and R code and (2) some real-world food sampling plans.	
The Guidance addresses a very important issue and makes some useful progress. However, the document needs to address several general issues, outlined below. More specific issues are made in the embedded comments within the document. CODEX as an organization, and CODEX Committee on Methods of Analysis and Sampling in particular, should produce guidance documents that are harmonized each with other. However, this is not the case between this document and others CODEX documents that already fully implement the concept of measurement uncertainty. These include: CODEX Guidelines on estimation of uncertainty of results, (CAC/GL 59-2006); Guidance on Measurement Uncertainty (prepared by UK to the 28th Committee Session in Budapest 2007); and the EU comments to Revision of the Guidelines on Measurement Uncertainty, prepared to the recent 40th Committee Session in Budapest 2019.	**EURACHEM**

The current document (CX/MAS 20/41/9) mainly uses the term 'measurement error' (ME), rather than 'measurement uncertainty' (MU). Although these two concepts and terms are quite different (as defined in VIM, 2008), the two terms seem to be used here almost interchangeably. The concept of MU is briefly discussed in the current document (e.g. Section 8.4), and usefully considered in the context of Compliance Assessment (Section 8.5), but the difference between ME and MU is not explained, nor the reason for using ME being used in preference to MU in most of the document. A further unexplained omission is the exclusion of the component of MU that arises from the sampling process. In 2007, an international collaboration between Eurachem, CITAC, Eurolab, Nordtest and RSC/AMC published the first edition of 'Measurement uncertainty arising from sampling. A guide to methods and approaches', and this has been recently updated[1]. This Guidance needs to be integrated into the methodology, and referenced, within your CODEX 'Methods of Analysis and Sampling', which will be thereby be significantly improved. The suggestion in Appendix III that 'that allowance for measurement uncertainty is not required' is clearly a retrograde and could lead to misclassification of food batches. The assertion that 'measurement error is generally insignificant compared to the variation between sample units' ('error' presumable meaning 'uncertainty') ignores evidence from a meta-analysis of measurement uncertainty for 67 food/analyte combinations over many different sectors[2], that reported average expanded measurement uncertainty as 57% and the analytical component of uncertainty alone is 21%. This level of measurement uncertainty is clearly not insignificant, as on average it accounts for around 33% of the total variance (which includes the between-batch variance). [1] Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches. Second Edition, Eurachem (2019). M H Ramsey, S L R Ellison and P Rostron (eds.) ISBN (978-0-948926-35-8). Available from http://www.eurachem.org [2] Ellison, SLR, Ramsey MH, Lawrance P, Stuart B, Minguez J, Walker MJ, (2017) Is measurement uncertainty from sampling related to analyte concentration?	
IUFoST supports the concept of reasonable sampling plans based on the commodity involved, cost of sampling and analysis, and reason for sampling be it health concerns, quality issues, or control of food hygiene and filth or other contamination, chemical or microbiological, in foods. General guidance for all of these factors is likely to be so general as to be meaningless, specific sampling plans within different food categories and different reasons for sampling are to be preferred.	**IUFOST**

SPECIFIC COMMENTS	**MEMBER / OBSERVER AND RATIONALE**
Purpose	
	Chile The General Guidelines on Sampling (CXG 50 - 2004) will be ~~referred to as~~ <u>called</u> 'the Guidelines' in this document.

	These Guidelines describe the design and evaluation of sampling plan for the international trade of food commodities -normally used term- or modify the text for international trade in food. Foods are frequently sampled throughout ~~in~~ the supply chain from producers to consumers, for the purposes of checking their quality. Clear definition of sampling plans is an integral part of specifications for foods. Sampling plans are included in Codex standards and may be used by governments in standards for foods. In Codex, sampling plans, in conjunction with methods of analysis, are intended ~~as a means of verifying~~ to verify that foods comply with ~~particular~~ (editorial) provisions such as composition, limits for contaminants and pesticide (editorial) residues and microbiological criteria that are part of Codex standards. Sampling therefore has an important role in achieving (editorial) the Codex objectives of protecting consumers' health and ensuring fair (editorial) practices in the food trade. Codex sampling plans also have a role in avoiding or removing difficulties which may ~~arise~~ (editorial) be created by diverging legal, administrative and technical approaches to sampling and by diverging interpretation of results of analysis in relation to lots or consignments of foods, in the light of the relevant provision(s) (editorial) of the applicable Codex standard. It is important that sampling is undertaken in a way that contributes to these objectives. A Codex standard (editorial) may set out a specific sampling plan for a particular context, or it may specify the outcome ~~to~~ that should (editorial) be achieved by a sampling plan. The main aim of sampling is to ensure that the consumer receives ~~product~~ food of acceptable quality. In addition, sampling plans should be designed to provide a high rate of acceptance of compliant ~~product~~ foods. The 'outcome to be achieved' (editorial) therefore defines allowable risks (the ones that may be permitted or there is a trend to permit them; they are not permitted or allowed risks) for the consumer and the producer. ~~Commodity~~ Food committees should define sampling plans ~~for~~ in accordance with the provisions in Codex standards. Codex methods of sampling should be designed in a way as to ensure that fair and valid sampling procedures are being used when food is being tested for compliance with a particular Codex food ~~commodity~~ standard. The design of these sampling plans should be based on the principles described in these Guidelines. These guidelines (editorial) provide recommendations on the design of sampling plans that achieve these aims.

A Codex standard can establish a specific sampling plan for a particular context, or it can specify the result to be achieved by a sampling plan. The main objective of sampling is to ensure that the consumer receives a product of acceptable quality. In addition, sampling plans must be designed to provide a high acceptance rate for conforming products.	**El Salvador** • In 1.0 Purpose of the Guidelines: include here the term of “safe” in the part of the text concerning sampling where the main objective is stated "... receives a safe product of acceptable quality". For reasons that the competent authorities of the different countries monitor the issue of food safety.
Basis of the Guidelines	
	USA §1, p4, line 15, “Sampling 4 therefore has an ..”; line 19, “… Codex standard 5.” Footnote 4 is not necessary. Footnote 5 as a style of citation is not efficient. If not mandated by Codex convention, do not use footnote as a means of citation. Adding a section of “References” in the end is much preferred. The whole paragraph between footnotes 4 and 5 reads awkwardly.
	Chile In ~~Throughout~~ these Guidelines, the two parties involved in a transaction in the food chain are referred to as the 'producer' and the 'consumer'. The terms 'producer' and 'consumer' are conventional and may apply to a range of different operators in the food chain, such as a grower, manufacturer, supplier, exporting country, processor, on-seller and customer or importing country. Sampling involves the selection of ~~small quantities~~ a specific number of food units from a lot allowing conclusions to be drawn about the lot from the results of the inspection or testing of the samples. Sampling reduces costs and prevents food loss through destructive testing, but it should be considered that despite applying a sampling procedure to the food lot under inspection, a probability of risks ~~inevitably creates risks~~ for both consumers and producers may exist. It is sometimes called "sampling inspection" (editorial) to distinguish it from the usual interpretation of sampling, which means the process of physically taking samples (editorial) from a lot. "Principles for the Establishment or Selection of Codex Sampling Procedures", Codex Procedure Manual. CX/MAS 20/41/9 These Guidelines (editorial) describe the inputs needed for the design of sampling plans, in order to accurately define the sampling situation, the acceptable levels and the acceptance conditions of a lot. This is covered in section (editorial) 3.

	The Guidelines (editorial) also provide information on the design and evaluation of sampling plans in sections 4 and 6, respectively. They will also help both consumers and producers design sampling plans that are appropriate for inspections of lots or consignments of food for acceptance according to the inputs as described. Their objective is to help the interpretation of the results of the food analysis by making a decision about the provision (acceptance or rejection) of the food when some characteristic of the food is subject to said provision. The Guidelines are based on the principles expressed in the (editorial) Principles for the Use of Sampling and (editorial) Analysis in International Food Trade (CXG 83-2013).
Throughout these Guidelines the two parties involved in a transaction in the food chain are referred to as the *'producer'* and the *'consumer'*. The terms 'producer' and 'consumer' are conventional and may apply to a range of different operators in the food chain, such as a grower, manufacturer, supplier, exporting country, processor, on-seller, and customer or importing country.	
	USA The term "food chain" should be replaced by "food supply chain". "Food chain" means differently in biology such as the food chain in ecosystem.
These Guidelines describe the inputs needed for the design of sampling plans, in order to accurately define the sampling situation, the acceptable levels of ~~a~~nd the acceptance conditions of a lot. This is covered in section 3.	
	USA "the acceptable level of" what?
The Guidelines are based on the principles expressed in the *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83-2013).	
	USA The reference to CXG 83-2013 can be entered in the "References" section. So the paragraph would read like this: The Guidelines are based on the principles expressed in CXG83-2013 (Ref. xx). This style of citation can save a lot of space in the main texts, where the title of CXG 83 is repeated in many places. For example, footnote 7 on p6 is not necessary. The Guidelines are based on the principles expressed in the Principles for the Use of Sampling and Testing in International Food Trade (CXG 83-2013).
Application of the Guidelines	
	USA 3§1.2, why need to specify that the sampling plans "might not directly address the producer's perspective." Should "fairness" require that the plans would be unbiased for either party? §1 should emphasize (using highlighted font or other means) several key concepts related to sampling plan: fairness, fitness for purpose, by probability, risks, producers, consumers and others. Once defined, these concepts do not need to be defined again (and again) in the other places of the "RG".
	Chile The Guidelines are intended primarily for use by Codex commodity committees responsible for developing sampling plans for provisions in Codex standards and by governments responsible for import or export

	inspection of foods. However the Guidelines are applicable quite generally and could (editorial) be used by any party engaged in the trade or sale of foods. For instance they could be used by any two parties at any stage of the supply chain, with appropriate consideration of the fairness of the transaction; or they could be used by a single party, for instance by a processor using a sampling plan for end-product verification. It should be noted that Codex sampling plans could (editorial) not directly address the producer's perspective. Producers should be (editorial) aware that the purpose of these ~~sampling~~ sampling plans is to specify the outcome to be achieved in terms of acceptable (editorial) risks. When a sampling plan ~~or an outcome~~ is clearly set out and a no-conforming outcome is obtained for a specific parameter, producers can design appropriate control procedures to achieve ~~them~~ conformity. The Guidelines also provide information on the evaluation of sampling plans obtained from other sources, to allow subjects (editorial) related to fairness to be investigated. 1.3 Codex commodity committees (editorial) In some situations, such as when measurement error is significant, it might not be possible to specify a standard (editorial) plan, suitable for general use. To ~~solve this~~ authorize the consumption of this food, commodity committees can specify criteria in terms of allowable risks that sampling plans are expected to achieve so that users can develop plans specific to their situation. However, commodity committees (editorial) should not just specify outcomes to be achieved without actually deriving (editorial) sampling plans and assessing their fitness for purpose (editorial) and the implications on producers in terms of fairness.
1.3 Codex Committee Committees In some situations, such as when measurement error is significant, it might not be possible to specify a standard plan, suitable for general use. To overcome this commodity committees can specify criteria in terms of allowable risks that sampling plans are expected to achieve so that users can develop plans specific to their situation.	
	EURACHEM Measurement error is not the same as measurement uncertainty (see VIM[2008] definitions). It should be made clear which terms is applicable to a given circumstance, throughout the document.
2. Concepts of sampling **2.1 Approach to sampling**	
	Chile 2.1 Sampling approach In the context of sampling, risk refers to the probability of making an incorrect decision about a lot of product, of either incorrectly accepting a lot of poor quality or of rejecting a lot of good quality product.

	There are three possible approaches to sampling: a. 100% inspection, involving inspection of all (i.e. 100%) of the product; (editorial) b. sampling based on the principles of probability; and(editorial) c. ad hoc inspection, that is, a sampling plan without a statistical basis. (editorial)
In the context of sampling, risk[1] refers to the probability of making an incorrect decision about a lot of product, of either incorrectly accepting a lot of poor quality or of rejecting a lot of good quality product.	
	USA §2.1 Footnote 6 is not necessary. Should Type I and II errors introduced here? Readers with statistics training would naturally think risk as these two types of errors. Indeed, much of the statistics in sampling plan is about Type I and II error calculation.
The risks and costs associated with each of these three options can be considered:	
	Chile The use of "risk" here relates to probabilities of incorrect acceptance or rejection of a product. This is different to the way that risk is usually understood in Codex, i.e. probability of an adverse health event; where the term refers to "public health risk".(editorial)
Approach (c)	
	Chile Regarding Approach (editorial) (a), 100% sampling, it is only possible to use this type of sampling when the test is not destructive and the volume of products is small. If the test is destructive, it is completely discarded and if the volume of products is large, it is impractical.
Approach (b)	
	Chile Approach (b) ~~has the disadvantage of higher risks as compared to approach (a), since some product will not be inspected. However by~~ using the probability approach the risks can be calculated and a sampling plan chosen that ensures these risks are controlled to desired levels, it ~~. It~~ also has the advantage of practicability and lower costs.
Approach (b), the probability approach, is described in detail later.	**USA** Line 4 from bottom of p5. “is described in detail later”, where?
Approach (c)	
Approach (c) is not recommended. …….. Decisions on acceptance or rejection should not be made solely on the basis of these ~~plans such a plan~~ plans	**USA** editorial

[31][1] The use of ‘risk’ here relates to probabilities of incorrect acceptance or rejection of product. This is different to the way that risk is usually understood in Codex, i.e. probability of an adverse health event; where the term refers to ‘public health risk’.

except by mutual agreement of the consumer and producer based on an understanding of those risks.	
	Chile Approach (c) is not recommended. It may be used for practical reasons, such as limited resources, or for simplicity. However such plans might not provide the expected level of assurance of food quality and may inadvertently impose high costs, for instance through unwarranted acceptance of food that could lead to illness or unwarranted rejection that in turn, could lead to the imposition of fines or penalties, trade sanctions or loss of access to markets. The risks associated with such plans should be evaluated where possible. Decisions on acceptance or rejection should not be made solely on the basis of these plans such a plan except by mutual agreement of the consumer and producer based on an understanding of those risks. The approach to sampling should be based on control of the levels of assurance provided and the costs to the parties involved in the transaction.
2.2. The probability approach **2.2.1 Control of risks**	
It is not possible to provide 100% assurance that all product in a lot complies with a specification when sampling is used. ~~There are two types of risks that can occur:~~	**Chile** It is not possible to provide 100% assurance that all product in a lot complies with a specification when sampling is used. Although it is not possible to provide 100% assurance that all product in a lot complies with a given specification, however, when using a sample under a standard (e.g. ISO 2859) for attributes inspection with an acceptable quality (AQL) by lot, a percentage of confidence can be guaranteed according to a statistical probability, for example with 95% confidence.
Sampling plans should be designed to control the risks to desired levels, i.e. they should take account of the principle of fitness for purpose[4]. Such control provides assurance, over the longer term, across many lots (i.e. in terms of probability).	**USA** 5. §2.2, "… across many lots (i.e. in terms of probability)". Is sampling plan providing control over just one lot?
Operating Characteristic (OC) curve:	
This is defined in ISO 3534 as:	**USA** There is no need to cite ISO 3534 for the OC definition.
	USA The OC curve introduced on p7 is far from easy understanding. At least, the graph needs a clearer graphic annotation of the areas for PR and CR. Seems that PRQ and CRQ could be explicitly stated as the lower

	and upper bounds of quality level, and one can relate PR and CR to the right and left tail areas, respectively, in a distribution curve (binomial, for example). I think a binomial example can be used to illustrate the above relationships
the Consumer's Risk Point (CRQ) is the level non-conforming in a lot corresponding to the Consumer's Risk.	**USA** Line 1 on p8. "Consumer's Risk Point"? or "Consumer's Risk Quality"?
Note that the choice of these points is arbitrary, but is conventional to define points on the curve corresponding to Producer's and Consumer's risks. The choice of Producer's and Consumer's risks, of 5% rejection and 10% acceptance respectively, is also arbitrary, but those values have also tended to become default values, so that designers of plans need to be concerned only with selecting appropriate levels for the Producer's and Consumer's Risk Quality.	**USA** Are 5% and 10% the commonly applied PR and CR, respectively, in food industry?
2.3.1 Random sampling	
	EURACHEM It is good that this topic is discussed, but there should also be a description of the two other basic modes, stratified and systematic sampling modes. If the heterogeneity type is not random (e.g. because of segregation) and the purpose is to estimate the mean and it uncertainty of the lot to be inspected, these two modes give much more reliable results. Consequently, the inspection cost can a fraction of that of a design based on assuming randomness. Further, stratified mode is never worse than the random plan, if the same number of primary samples are taken.
It is common for lots to be "layered", individual items might (say) be packed in cartons, there might be several (but the same number) of these smaller cartons packed into a larger carton, and several (but the same number) of the larger cartons packed on a pallet. Selecting a random sample of size "n" items would proceed as follows:	**Chile** Add "as per the used standard:" at the end of the paragraph. It is common for lots to be "layered", individual items might (say) be packed in cartons, there might be several (but the same number) of these smaller cartons packed into a larger carton, and several (but the same number) of the larger cartons packed on a pallet. Selecting a random sample of size "n" items would proceed as follows, as per the used standard:

select 'n' ~~pellets~~ pallets from the number of pallets in the lot;	**Mauritius and CCTA**
For bulk materials taking a random sample is more difficult. Many lots of bulk materials can be considered as a collection of segments; segments are selected at random from the total number of segments, then within each segment that has been chosen a random sample of increments is taken.	**USA** §2.3.1 "sample of increments" – not a good choice of words.
In principle there is no need for random sampling for well-mixed fluids or bulk products; however random sampling might still be used as a precaution against inhomogeneity or for procedural reasons.	**Chile** In principle there is no need for random sampling for well-mixed fluids or bulk products; however random sampling might still be used as a precaution against inhomogeneity or for procedural reasons. For the above cases, it is important to consider that in order to carry out random sampling correctly, it is essential to use instruments and types of samples appropriate for this purpose.
3.1.1. Introduction	
	Chile Make it clear that the statistical inputs may consider more parameters than those listed here, depending on the destination of the product. Rationale: Statistical inputs
There are two types of inputs to the design of sampling plans; **administrative inputs** such as descriptions of the food or Codex provision the sampling plan applies to and **statistical inputs**, which are those required for the design of the sampling plans.	**CCTA**
3.1.2 administrative inputs to the design of sampling plans	
	USA Is "provision" synonymous to "specification"? Since "spec" is often used in Quality literature, a clear distinction between these two words is required. If they are the same, then use only one.
	USA "Attribute" in the table. Here, "attribute" seems to mean "trait" or "factor", of both continuous and discrete nature. Its meaning here is different from that in "attribute plan", where "attribute" means only the discrete

	factors (traits). Its meaning is different from the same word in the table below on the same page.
3.1.3 Generic statistical inputs to the design of sampling plans	
	USA "whether the specification applies to every item in a lot*…", and the endnote * to the table. Delete them, because 100% test is out of the scope covered by sampling plan. In the same text box, "the acceptance conditions of a lot controlled", delete "controlled" because it is unclear what this term mean in this context. In the table under §3.1.3, "Composition of the lot*" – no endnote for *. On the same row, 'Refer to the example below' –unclear which example this sentence refers to (if not use the hyperlink).
	USA PRQ/AQL, CRQ/LQL – choose only one set of terms.
Example stringency	
	Mauritius in case of certain micro-organisms, would the figure not be 0% instead of 1%?
4. Design of sampling plans	
The Codex Procedure Manual and the *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83-2013) state that Codex methods of sampling should be designed to ensure that *'fair and valid sampling procedures are used when food is being tested for compliance with a particular Codex commodity standard'*.	**Chile** The paragraph should not be allowed to be less strict than in food safety since the consumer could have stricter criteria for sampling plans and these ones should be considered. This applies mainly in export cases. The term "cierta" in the Spanish translation (for certain) es a rather vague term that does not reflect the intent of the English text. The term "certain" in the English text refers to information that is true or veracious.
The Procedure Manual recommends that when commodity committees have included provisions on sampling in a Codex commodity standard, these should be referred to the Codex Committee of Methods of Analysis and Sampling for endorsement along with certain information relating to the sampling plan.	**Chile** The first part of the paragraph may lead to confusion (it is suggested to eliminate it). The second part of the paragraph could substitute the previous paragraph on the relativity among parameters, or the two might be consolidated as to avoid repetition. I suggest to consolidate the two paragraphs and improve their wording since there are duplications en the two paragraphs.
4.1 *Design of sampling plans for various situations*	

	USA The sentence “Sampling plans for each of these situations can be derived from apps” – against designing a sampling plans only based on using apps, without first understanding the underlying statistical theory. Apps as a tool provides computation convenience, but the plans must be based on concrete statistic theories or formulas. And the readers should be able to easily modify any apps to suit their own unique requirements. Given the universal literacy of R among QC workers, an open source of the Apps R code would be a welcome addition to “RG”.
This section provides information on the types of sampling plans that are appropriate for different situations. Sampling plans for each of these situations can be derived from apps or the statistical literature, including standards in some cases. The remainder of this ‘Design of Sampling Plans’ section discusses the design of sampling plans for various situations, depending on the nature of the lot under inspection, the nature of the measurements made, the presence of significant measurement error and other considerations.	**Mauritius** is it possible to provide more details on the apps?
4.1.1 Roadmap	
	USA against the use of ISO standards. We can consult ISO standards for some inspiration (for example, using ISO standards as a benchmark to “RG”), but we should not expect that our targeted readers to consult ISO standards when using Codex.
	USA Also, seems that in the column headed by “CXG 50 reference”, the second and third 4.2.3. for “General Design” and “FNC” should be “4.4.3”.
	Chile MEASUREMENT ERROR Type Error Negligible Significant Plan Type Attributes Variables Attributes Variables Lot Type Discrete Lots • ISO 2859 /CXG 50 point 9 • General Design/ CXG 50 4.2.3.• ISO 3951/ CXG 50 point 4.3.3. • General Design/ CXG 50 point 4.2.3. ----- • ISO3951-1 Annex O/ CXG 50 point 4.3.3.

	• ISO 3951-6/ CXG 50 point 4.3.3. • Repeatability adjustment/ CXG 50 point 4.3.3. Fractional Non-Conformance (FNC)/ CXG 50 point 4.2.3. Bulk Materials --------- • Beta distributed parameters: CXG 50 point 4.4.1 & 4.4.3 • Compositional proportions: CXG 50 point • ISO 3951-1 Annex O/ CXG 50 point 4.4.2. • ISO 3951-6/ CXG 50 point 4.4.2. • Repeatability adjustment • FNC/ CXG 50 point 4.4.2.
Flow diagram	
	Chile The Spanish version should include a chart translated to Spanish.
4.2 Sampling plans for inspection by attributes **4.2.1 Introduction**	
	USA §4.2.1 last sentence on p12 "… are measured on a scale." Unclear the meaning. Does this sentence mean that some categorical attributes can be ordinal? (Their order on a scale is meaningful, as compared to some attributes where the order does not matter).
4.2.2 Forms of attributes sampling plans	
	USA P13, line 6. "…found in the sample". "Found" should be "allowed"? on the same page, second paragraph, the use of "economy" is obscure. "Some economy in the number of samples" seems to mean that "the population (lot) size" must be taken account for.
Zero-acceptance number (ZAN) plans	
ZAN plans are a special case of two-class plans in which the acceptance numbers are set to c=0. They are used in more critical situations where only consumer's risk is considered directly. These plans	**USA** "Zero-acceptance number (ZAN) plans". Delete "metrological applications" – seems that metrology has no direct application in food industry. "Three-class attribute plans" When "none of the n samples has a level exceeding M", does the 3-class plans shrink to 2-class (binomial)

are often used in ~~metrological~~ applications or situations such as pathogens or for foreign matter where acceptance of lots demands that non-conforming items are not found in the inspection.	plan?
Refer to the inputs section for baseline information.	**USA** Where is “baseline information”?
4.3 Sampling plans for inspection by variables	
	Chile Define the term "variable" to give the concept a consistent meaning.
4.3.1 Introduction	
These plans are usually referred to as **Variables Sampling Plans**. If the underlying distribution of individual measurements is known, acceptance sampling can be performed directly on the measurements themselves. This often allows a considerable saving in sample size but we need to know the probability distribution of the underlying measurements. The Gaussian or normal distribution is commonly adopted as the distribution of the measurements. For compositional proportions in bulk material, the beta distribution is more appropriate but the normal distribution can serve as an approximation.	**Canada** Given the controversy over the variables sampling, perhaps the discussion could remain in the main section, but include a listing of caveats, including those mentioned by the USA. Perhaps specific details about more complex sampling plans (e.g. beta distributions) could be excluded and included in a separate technical document.
• the extent of conformity of each unit is taken into account in the application of the plan; and	**Canada** As variables sampling treats variables as continuous and does not categorize them, it has greater power and thus does not require as large a sample size as attributes plans. They also state the distribution only needs to be a reasonable approximation to the data to provide legitimate results, which is generally true for most parametric analysis. One of the listed advantages of the variable sampling method is that it can incorporate measurement error directly into the sampling design. Deleting measurement error from consideration in the sampling design is recommended by Canada, although it is recognized that this removes one of the practical advantages to using variable sampling.

4.3.2 Form of variables sampling plans	
	USA Elaboration of k is needed here. Naïve readers may simply think k is a quantity similar to z (as applied to standard normal) or t (from t- distribution when n is limited), and $\bar{X} \pm kS$ is the C.I. of sample mean. How k is related to z or t (and n), and how k is measured against Type I and Type II errors, should be discussed or referenced. In the formulas, $\bar{X}+kS \leq U$ and $\bar{X}-kS \geq L$, the meaning of S is not clear. Does S stands for s.d. of individual X, or it stands for s.e. of the sample mean ($\bar{X}$)?
In variables plans, the mean ~~(~~(x), is compared with the acceptance limit in a similar way to the attributes plans but, in order to allow for the variability in the lot, the sample standard deviation 'S' is computed.	**CCTA** please include the correct symbol
Negligible measurement error	
	Canada Suggest removing measurement error throughout the document as this is considered in other documents and not generally as a part of food commodity sampling plans
4.4 Sampling of bulk materials	
	Chile Bulk materials are continuous, consisting for example of particles of different density and sizes etc. They may be in liquid or solid form, such as food in powder. ~~An example is milk powder.~~ It is impossible to view bulk materials present in a lot as a set of distinct objects because there is no way of selecting the items one by one in a way that is not biased when using simple random sampling. (editorial) This is where a different methodology is introduced, which brings with it sampling bias and non-representativeness.
4.4.1 Introduction	
	USA The first paragraph on p16 seems to imply that bulk materials cannot be randomly sampled without bias and representativeness. That is not correct. Bulk materials can be randomly sampled. The bullets in the second paragraph of §4.4.1 need to be revised because some of them are not uniquely only applicable to bulk materials. For example: • Acceptance on a lot-by-lot basis. That applies to all sampling plans. • Characterize the materials as to grade. Grading is not limited to bulk material. • Weight or content determination. Not limited to bulk.

	Chile Provide examples on "some kind" of sampling device. (May be a translation error that does not reflect the intent of the paragraph).
Bulk materials are continuous, consisting for example of particles of different density and sizes etc. An example is milk powder. It is impossible to view bulk materials present in a lot as a set of distinct objects because there is no way of selecting the items one by one in a way that is not biased when using simple random sampling. This is where a different methodology is introduced, which brings with it sampling bias and non-representativeness.	**EURACHEM** There needs to be some discussion of how to minimize sampling bias, by design of both sampling equipment and the sampling plans. Sampling bias is at its lowest when sampling moving bulk materials. Because such sampling bias is a transient and ever changing phenomenon, there is no way to correct for it afterwards. (Minkkinen P (2004) Practical applications of sampling theory. Chemometrics and Intelligent Lab. Systems, 74, 85–94)
Bulk materials being continuous means parts of samples can be mixed together to form a composite. This composite then gets tested only once, rather than having to do many tests on the individual parts. This is a physical way of creating a composite sample representing the average content of lot.	**CCTA, EURACHEM, Chile** Text repeated from previous paragraph - so delete
Refer to the inputs section for baseline information.	**Chile** The indication to refer to another section should state with more clarity what input to consult.
4.4.3 Sampling plans for compositional proportions (where measurement error is negligible)	
	USA The formula after the first paragraph on p18. Delete the parenthesis. Does the parameter k defined in the same way (using the same table of k values) as before in the variable sampling plans (p15, §4.3.2)? When P is close to either 0 or 1, the Ward-interval type of approximation may not be a good method.
Special sampling plans	**Chile** Indicates that it is in progress, awaiting information to comment on.
	USA 17. §5. Information box under §5.

	"A minimum of the PRQ, CRQ and the associated risks" – which quantities, PRQ/PR, CRQ/CR, can be minimized? Last paragraph of §5, p18. "In the case details of the …along with details of the sampling plan itself." – which plan? The one submitted to CCMAS or the one used in real-world application?
reference to the source of the sampling plan;	**Chile** It is suggested to change for: reference of the sampling plan to a standard or bibliography *Category : SUBSTANTIVE*
Of course the outputs from the design process differ from those needed in a submission to CCMAS or other parties for that sampling plan to be considered for approval. In this case details of the inputs and possibly some justification for the choice of values for those inputs would also be required along with details of the sampling plan itself. There should also be some consideration of the impact of the proposed plan on producers, both in terms of rejection of product and the plans producers themselves might have to use.	**USA** "In the case details of the …along with details of the sampling plan itself." – which plan? The one submitted to CCMAS or the one used in real-world application?
In the general sense, 'fitness for purpose' entails cost, practicality and fairness.	**Mauritius** the definition "fitness for purpose" should also refer about the technical adequacy as the term "practicality" has different meanings.
Practicality	**USA** §6.1.2. p19, the two paragraphs, starting with the sentence "The designer of sampling…" and ending with the sentence "… and other costs can also be taken into account", should be combined into one and cut the repetitiveness. It is hard to follow the rationale behind 'indifference' plans.
Practicality	**El Salvador** • In 6.1.2 "Practicality" it is suggested to to change the subtitle for "Cost and Practicality" since the text deals with both subjects, and it is likewise mentioned in 6.1.1. referring to "Cost".
Sampling and testing procedures are fit for purpose in a given product assessment, if when used in conjunction with appropriate decision criteria, they have acceptable probabilities of wrongly	**Chile** Suggestion is made to substitute for the following text: Sampling and analysis procedures are fit for purpose in a given product assessment if, when used in conjunction with appropriate decision criteria, they have acceptable probabilities of wrongly accepting or wrongly rejecting a lot or consignment.

accepting or wrongly rejecting a lot or consignment.	*Category : SUBSTANTIVE*
Fairness	**El Salvador** • In 6.1.3 Fairness; reference is made to the CXG 83 standard, it is suggested to add the name of the guideline "Principles for the use of sampling and analysis in international trade", providing the name of the regulation referred to in the text. *Category : EDITORIAL*
Sampling plans having inappropriate stringency, not commensurate with the situation e.g. plans for assessment of composition that are more stringent than those for food safety.	**Chile** Editorial suggestion to use another synonym of fairness in Spanish. It is suggested to include, as an example, in the text on sampling with inadequate strictness: also the possibility of food safety plans that are not strict enough. Plans not based on statistically not valid principles. *Category : SUBSTANTIVE*
Physical sampling procedures	**El Salvador** • In 7.1 "Physical sampling procedure" we will provide our comments when the proposed text is available. *Category : TECHNICAL*
Physical sampling procedures	**Chile** It is hard to express comments since it is a text in course of development. Sampling is presented only for convenience and no details are provided for other stricter types of sampling. There are ambiguous paragraphs.
Physical sampling procedures	**USA** §7.1.1 The bottom sentence of p20. This is a good example of wordy sentence. Should be rewritten to something as "… use of convenience sampling might be a procedure neither fair nor valid."
Often the first indication of possible issues with sampling and testing occurs when there is a dispute over the assessment of product. Disputes between parties in trade may occur for many reasons including differences in the testing between the laboratories concerned; the existence, appropriateness and statistical validity of the sampling plan used to assess the product; the allowances made for general	**USA** §7.2. First paragraph. Delete "product variation".

measurement error and within-lot variation ~~product variation~~; differences in physical sampling procedures; differences in composition of the samples tested due to product inhomogeneity or changes occurring during storage and/or transport of the product.	
Disputed lots provide motivation for the development of sampling plans which, in accordance with CXG 83, should have been developed before trade commenced. The same process should be followed in the design of the sampling plan; within-lot product variation and measurement error would be standard inputs in to the design of a sampling plan.	**USA** §7.2.1. The second paragraph under this sub-section. "The same process should …to the design of a sampling plan." Does this sentence mean that the sampling plan from both parties (producer and customer) should be the same?
Re-inspection	**Chile** The possibility of accepting unacceptable lots grows. Furthermore it is not a recommendable practice in microbiology due to the heterogeneous distribution of pathogens. *Category : SUBSTANTIVE*
When the original inspection results are suspect, the provision for lot re-inspection can be incorporated. Re-inspection is done when a lot was rejected on the first inspection, but the lot is resubmitted for acceptance inspection so that a new sample can be taken to make a decision. This process can be repeated but the design of the sampling plan depending on the number of re-inspections allowed.	**Canada** It was stated that re-inspection of lots was statistically invalid. Please clarify why this is the case. Are there statistical assumptions violated, or is it invalid because it does not preserve the level of consumer's risk (i.e. it increases it)? The methodology corresponds to a published article from Govindaraju and Ganesalingam (1997). This procedure increases consumer's risk by a non-zero amount, which may not be appropriate in many situations. However, this method does control for a specified "acceptable" level of consumer's risk using the specified statistical parameters for consumer's risk during sampling design phase. Instead of removing any mention of re-inspection, is it possible to state the limitations/pitfalls of such a plan, and explain situations where re-inspection plans are invalid and why. *Category : SUBSTANTIVE*
Inhomogeneous lots might occur because inspection lots differ from manufacturing lots or other reasons.	**USA** §7.2.3. "inspection lots differ from manufacturing lots" – should inspection by default is lot-by-lot; therefore, manufacturing lot and inspection lot must be the same. I think what the first sentence under §7.2.3 simply means there are heterogeneity within a lot so sample is not representative. The whole subsection talks about heterogeneity (within lots and among lots) that will decrease the power to sampling inspection. A solution would be either to pool or subdivide the lots. This subsection was well written but some wordings were just confusing.

Lot heterogeneity usually increases producer's and consumer's risks, so that consumer protection may be compromised when an inspection lot is not homogeneous.	**EURACHEM** There needs to be more discussion of how to allow for different types of heterogeneity (e.g. random and segregated) in the sampling target (e.g. batch of material). It is particularly important as heterogeneity is the ultimate source of measurement uncertainty arising from sampling, and hence of potential misclassification of the material. *Category : SUBSTANTIVE*
Other Sampling Plans	**Chile** The sampling plans described in paragraph 8 don't add to the good interpretation of the concepts for their practical implementation. *Category : SUBSTANTIVE*
Other Sampling Plans	**USA** §8. "Compliance of the average level' is not included. This is rather surprising because, for example, it is totally conceivable that multiple samples drawn from a lot and measured for protein, and then the sample mean should pass certain cutoff for compliance inspection. Is this a 'compliance of the average level' type of sampling plan?
Microbiological sampling plans	**Mauritius** Section 8.1 on microbiological sampling plans is relevant but it should include all possible scenario e.g. 3 class plans as suggested.
Microbiological sampling plans	**El Salvador** • In 8.1 "Microbiological sampling plans" consider to include it in the guideline due to the importance of the subject. *Category : SUBSTANTIVE*
Microbiological sampling plans	**Chile** I think this sampling plan is indirectly explained in section 4. *Category : SUBSTANTIVE*
Microbiological sampling plans	**USA** §8.1. If GL 50 works fine for microbiological sampling, I don't see the point to include it in 'RG'.
[For review as to whether this stays in the Guidelines]	**Mauritius** Section 8.2 could be removed if not covered in Codex Provisions
Form of the sampling plan	**USA** §8.3.2. Delete one "either" in the only sentence under this subsection, and rewrite the sentence as "… based on inspection either by attribute or variable plans."
Conformity testing	**Australia** To help make this document more focused and concise, we would suggest section deletions including; As 'Conformity testing' or evaluation of conformity, is not mentioned in the 'New Work to revise the Guidelines on Sampling' we consider it out-of-scope for this guidance. Thus section 8.4 should be deleted.
Conformity testing	**USA** §8.4. I doubt conformity testing can find wide application (if any) in food industry. So 'conformity plan' should

	be excluded from “RG”.
[For review as to whether this stays in the Guidelines]	**Mauritius** many users of Codex standards are involved in conformity assessment activities, for example regulatory bodies. We are therefore of the opinion that the section should be maintained.
Measurement and sampling uncertainties, including metrological traceability, become crucial for the declaration of conformity, especially when the measured value is close to the set limiting value.	**EURACHEM** The uncertainty is in the measurement value, and arises not just from the analytical process, but also from the primary sampling process (Eurachem, 2019). Better wording would therefore be 'Measurement uncertainty arising from both chemical analysis and sampling, and metrological traceability, become crucial...'. Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches. Second Edition, Eurachem (2019). M H Ramsey, S L R Ellison and P Rostron (eds.). ISBN (978-0-948926-35-8). Available from http://www.eurachem.org *Category : TECHNICAL*
The main disadvantage of the conformity testing procedure is that in many cases, inconclusive results will be obtained even though a sample is conforming but due to measurement errors, the uncertainty interval includes the limiting value.	**EURACHEM** The term' measurement error' been used, but the item under discussion is 'measurement uncertainty' - so that would be the more appropriate term.
The ISO 10576 Standard does not encourage reduction of measurement errors by design and hence poorer measurement systems will produce more inconclusive results. Hence producers may be forced guard-band in order to reduce the incidence of inconclusive results from measurements.	**EURACHEM** Doesn't make sense. Should wording be 'Hence producers may be forced use a guard-band
General information	**USA** §9 General Information. I think this section is misplaced. It should be combined with §2 and provide a better explanation of “OC”. The table and figure in this section is good, but it would be better if the resolution in the PR area had a higher resolution in the graph. The readers have no way to tell PR’s among different sample sizes.
Normal inspection plan is the plan used when the process is considered to be operating at, or slightly better than, the Acceptance Quality Level (AQL).	**Canada** There are still inconsistencies with different terms for the same concept in the paper and the included apps to describe the same concept (e.g. PRQ and AQL).
Guidance for producers	**Australia** To help make this document more focused and concise, we would suggest section deletions including;

	The document needs to be balanced when explaining the guidance from a 'producer's' or 'consumer's' perspective. So a section with the heading 'Section 9.2 Guidance for producers' does not help in this. We suggest this section is deleted.
Guidance for producers	**El Salvador** • In 9.2 "Guidance for producers" consider it to be included in Guideline CXG50.
[For review as to whether this stays in the Guidelines]	**Mauritius** a good point is raised and we consider it relevant to be maintained in the draft.
Appendix: Definitions	**Egypt** Definitions: Egypt recommends adding a definition of "Decision Rule" and "Guard Band" as follows: - Decision Rule: a documented rule that describes how measurement uncertainty will be allocated with regard to accepting or rejecting an item "product" according to its specification and the result of a measurement. - Guard Band: interval between a tolerance limit and a corresponding acceptance limit. NOTE: The guard band includes the limits. N.B. Explanation of decision rules and guard band can be found in ISO 10576-1[1] and JCGM 106:2012[2]. [1] ISO 10576-1:2003, Statistical methods -- Guidelines for the evaluation of conformity with specified requirements -- Part 1: General principles. [2] JCGM 106:2012, Evaluation of measurement data – The role of measurement uncertainty in conformity assessment.
Appendix: Definitions	**EURACHEM** The text contains many abbreviations and symbols, and reference to normative documents without indicating their names, which complicates the process of understanding and using the document. It is recommended that two sections on "Abbreviations" and "References" be added (maybe as appendices)
Appendix: Definitions	**USA** §10. Appendix: Definitions. This section should be expanded to include acronyms used in "RG". "on-seller", what does that mean? Page 27. The two Information Notes of "representative samples" and "Confidence". Delete them because they only introduce confusing.
Acceptance sampling plan - plan which states the sample size(s) (1.2.26) to be used and the associated criteria for lot (1.2.4) acceptance.	**EURACHEM** Is this a reference to the relevant section of this standard? If so, please clarify

Draft supplementary e-book	**El Salvador** • El Salvador considers that the document in Appendix II Draft complementary e-Book presents a good sampling plan tool with links and applications to facilitate the use of CXG50. Suggest that direct access to the document be established and placed on the Codex website and that all members have on-line access without restriction. It is also suggested to link the access in guideline CXG 50. *Category : SUBSTANTIVE*
Producer's Risk (PR) Producer's Risk is the probability of wrongly rejecting a lot that is of acceptable quality. It is a point on the OC curve corresponding to a predetermined and usually high probability of acceptance.	**EURACHEM** Repeated text - so delete
App15 enables the user to design a variables sampling inspection plan that is adjusted for the repeatability SD of measurement errors. This app particularly shows that the acceptability constant *k* constant must be smaller depending on the size of the repeatability SD.	**EURACHEM** There is no App 14. Was the omitted on purpose?
The text by Schilling and Neubauer (2008) may be consulted for more details on the administration of acceptance sampling.	**EURACHEM** Please make explicit where to find the full references for this and other citations.
The first four purposes are particularly critical. The designed sampling plan must explicitly quantify the producer's and consumer's risks. Some of the published sampling inspection procedures such as those in ISO 3951 place more emphasis on reducing the producer's risk with increasing lot size. This is to encourage large scale production and lot formation. For international trade and particularly for food products, consumer's risk control is particularly important in addition to simplicity of operation and transparency and fairness in reducing the risks to both producers and consumers.	**Canada** It has been proposed that a change to the guidelines that obtains the desired sample size by simultaneously fixing both consumer and producer's risk at the desired level. Although this approach is more in line with conventional statistical practice in obtaining an adequate sample size, compared to a trial-and-error approach that examines only CR. However, as they involve changing two parameters (PR and CR), they are more complex to present, difficult to understand, and easier to misapply.

In this section, some of the commonly encountered issues such as the relationship between sample size and lot size are discussed. Resampling and retesting (not the same thing) are also discussed. Resampling is used to reduce the producer's risk when random sampling of the lot is difficult whereas retesting is a way of overcoming inaccuracy of test results due to measurement uncertainty. If measurement errors are expected to dominate, sampling inspection plans can be adjusted for measurement errors. This adjustment can be done fairly to protect both the producers and consumers. This topic is covered in detail in a later section (Section 9).	**EURACHEM** This section again uses the two terms 'measurement error' and 'measurement uncertainty' apparently interchangeably when they have different meanings. In most cases the term 'measurement uncertainty' is what is meant by 'measurement error', and replacements should be made.
The term measurement errors relates to numerical measurements on the quality characteristic of interest. The following definitions relating to measurement errors are based on ISO-5725.1 (1994).	**EURACHEM** Very old document, that was written before the concept of Measurement Uncertainty as the central metric for measurement quality (see VIM, 2008). Use of terminology, and its relationship to 'measurement uncertainty', therefore need to be included.
Error is the difference between the measured value and the true value of what is being measured. Errors can be either random or systematic. Random errors are uncorrelated, but they affect the results of the repeated measurements. Some examples are: whether they are repeatable, whether they are reproducible, and whether they are stable. Systematic errors are different, in that they affect all measurements taken in the same way and can be identified when the random errors are small. Some examples are: accuracy, bias, and drift.	**EURACHEM** Text should be added to this section to explain the difference between the error of a measurement and its uncertainty, using VIM(2008)
In order to adjust for the bias, the actual measurements can be converted to bias	**EURACHEM** This procedure assumes that the measurement bias (analytical and sampling) is known and quantified (with

adjusted measurements and then the variables plan can be applied.	negligible uncertainty). This limitation should be made explicit.
An observed measurement *Y* is classified with certainty as conforming or not for given specification limits only when there are no measurement errors. Analytical testing of fat content etc. involves considerable measurement uncertainty, often up to half of the observed variation. The distribution of the measurement errors (*Z*) can be fairly well ascertained using past calibration studies. Measurement error uncertainty results only in an estimated probability of conformance of a unit. The probability of nonconformance of an individual unit based on the error-prone measurement is defined as the *fractional* nonconforming unit. The following figure illustrates the concept of fractional nonconformance. Given the measurement error distribution, the probability of breaching the upper specification limit,  is the FNC value.	**EURACHEM** 'Measurement error uncertainty' is ambiguous, and not a term recognised by VIM(2008). Was the intended meaning really 'Measurement uncertainty produces only in an estimated probability of conformance of a unit.' ? or 'The wide confidence interval on the estimate of the measurement uncertainty produces only in an estimated probability of conformance of a unit.'?
The USA proposal for a general design of the revised Guidelines (Appendix III):	**Mexico** Mexico supports the US comment regarding continued work on updating the CXG50 and that the description of sampling in the document is complicated for non-experts. The inclusion of spreadsheets is also supported for a better understanding of sampling plans.
The comments provided by the USA to the EWG including the draft top-level outline for revised Guidelines	**Egypt** Egypt supports the comment received from the US regarding the revision should avoid examples.
Annex 1. Model Outline (CXG50)	**USA** **Preamble** (Brief aspects of revision, e.g. understandable to audience with limited statistical training) **Key terms and definitions** (Brief definitions) **1. Introduction** a. Prerequisite Codex documents (Procedural Manual, GCX 83) b. Target audience: Codex committees, governments, industry

c. Scope
   i. Food hazards, suitability for consumption, quality
   ii. Border inspection and other receiver-oriented situations
   iii. Receiving finished products or raw materials
   iv. Does not cover statistical process control (SPC)
   v. Does not cover multiple, sequential, or switching sampling schemes

2. Basic concepts

a. Reasons for sampling
   i. HACCP and GMPs control food safety, suitability, and quality
   ii. End-product testing is ineffective and discouraged
   iii. Acceptance sampling
      1. Acceptability of lots with unknown control history
      2. Used when no better tools are available (e.g. no HACCP documentation)
      3. Sample size limitations generally preclude reliable assessment of lot performance criteria.
      4. Routine (lot-by-lot) sampling, or intermittent (between lot) sampling
   iv. Surveillance sampling
      1. Assess performance of supplier control system (GMP/HACCP in place)
      2. Periodic lot sampling with more rigorous plan (more samples, lower action level)
      3. Follow-up when expected system performance is not met

b. Performance criteria based on food safety/quality objectives
   i. Fraction or percentage of non-conforming units that determines an unacceptable lot
      1. Health hazards (chemical, physical, biological); typical criteria: 1:1,000 to 1:1,000,000
      2. Suitability for human consumption (nutritional composition, filth, decomposition); similar criteria to health hazards
      3. Quality defects (e.g. color, texture, size, workmanship defects); typical criteria: 1:100 to “Six Sigma” (3.4 defects per million)
   ii. Definition of a non-conforming sample unit
      1. Maximum/minimum level (continuous variable, count, proportion)
      2. Presence/absence
   iii. Sample unit amount (weight, volume, count) (“decision unit”)
      1. Amount affects observed variance (heterogeneity, ‘averaging out’)
      2. Amount is based on risk assessment
         a. Acute health hazard (e.g. single serving)
         b. Chronic health hazard (e.g. average in larger amount)
         c. Quality defect (user awareness, e.g. retail package)

c. Probability and plan performance
   i. Sampling estimates a lot parameter based on a subset
   ii. Can calculate probability of accepting an unacceptable lot (based on performance criteria and sample size)
   iii. Can calculate the probability of rejecting an acceptable lot (based on an allowed

	fraction of non-compliant units) iv. Operating characteristic curve (OC curve) plots the probability of rejecting (or accepting) a lot versus the fraction (or percentage) of non-compliant units in the lot. 1. 'Consumer risk-point' 2. 'Producer risk-point' d. What is a lot? (see FAO/WHO 2016, Part 2) i. How defined ii. Why not redefined after sampling iii. Why invalid to re-sample **3. Two-class attribute sampling plans** a. Widespread use, distribution-free, universally applicable, simple, reliable i. Applicable to lots containing product from different production lines, dates, and raw material sources ii. Applicable to economic fraud (mixed quality) b. Binomial distribution i. Characteristic present or absent in sample unit ii. Characteristic under or over the limit in sample unit c. Zero-acceptance number plans (ZAN) i. Applicable in most situations; efficient ii. Design 1. Sample size not limiting (use required number of samples) a. OC curve passes through 'consumer's risk point' b. 'Consumer's risk point' – fraction of non-compliant units that should be rejected most of the time (e.g. 95%) 2. Sample size limiting (use less than required number of samples) a. Typically the case, e.g. 299 samples required to reject lots with 1:100 non-compliant units (with 95% probability) b. OC curve does not pass through the 'consumer's risk point' c. Lower probability (< 95%) of rejecting lot that should be rejected most of the time 3. 'Producer's risk point' a. Level of non-compliant units that should be accepted most of the time (e.g. 95%) due to limitation of industry control systems, or food security b. Generally met by default when sample size is limiting iii. General formula, hypergeometric formula, table, app link d. Two-class attribute plans using acceptance numbers i. Use when 'producer's risk point' is not met using ZAN plan ii. Increases discrimination, steepens OC curve iii. Requires many more samples to maintain 'consumer's risk point' iv. Formula, table (c = 1,2), app link **4. Situational sampling plans** a. Three-class attribute sampling plans

	i. Trinomial ('compliant units', 'non-compliant units', 'marginal units') ii. 'marginal units' – indicate inadequate GMPs/HACCP; still comply with health limit iii. App link b. Variables sampling plans (based on known distribution) i. Require fewer samples (presuming conditions are met) ii. Decision limit is proportional to the standard deviation (known or estimated) iii. Lots may be rejected when all sample units are compliant (can cause confusion) iv. Variables plans for fixed standard deviation (SD) 1. Used when SD known and stable 2. Within production facility or analogous situation 3. Continuous series, one facility, uniform raw material v. Variables plans for unknown standard deviation 1. Appropriate for screening trade lots (varying SD) 2. Unacceptable lots generally have higher SD 3. SD estimated from results of lot sampled 4. Different decision limit used for each lot vi. Plans for provisions based on the average in the lot vii. App links **5. Random sampling** a. Methods for packaged goods b. Methods for bulk materials **6. Composite samples** a. Use with attributes sampling plans i. Reduces analytical cost when lots are regularly acceptable ii. Analytical method suitable for reduced decision limit iii. Importance of composite homogeneity b. Use with variables sampling plans i. Reduce analytical cost ii. Importance of composite homogeneity **7. Sample handling** a. Drawing analytical sub-samples and importance of homogeneity b. Holding conditions (environment, time) c. Traceability **8. References** a. FAO/WHO Statistical Aspects of Microbiological Criteria Related to Foods b. FAO/WHO Public Health Risks of Histamine and other Biogenic Amines from Fish and Fishery Products c. Recommended Methods of Sampling for the Determination of Pesticide Residues for Compliance with MRLS CAC/GL 33-1999
Annex 2. Model Text, Basic Concepts of Sampling (CXG50)	**USA**

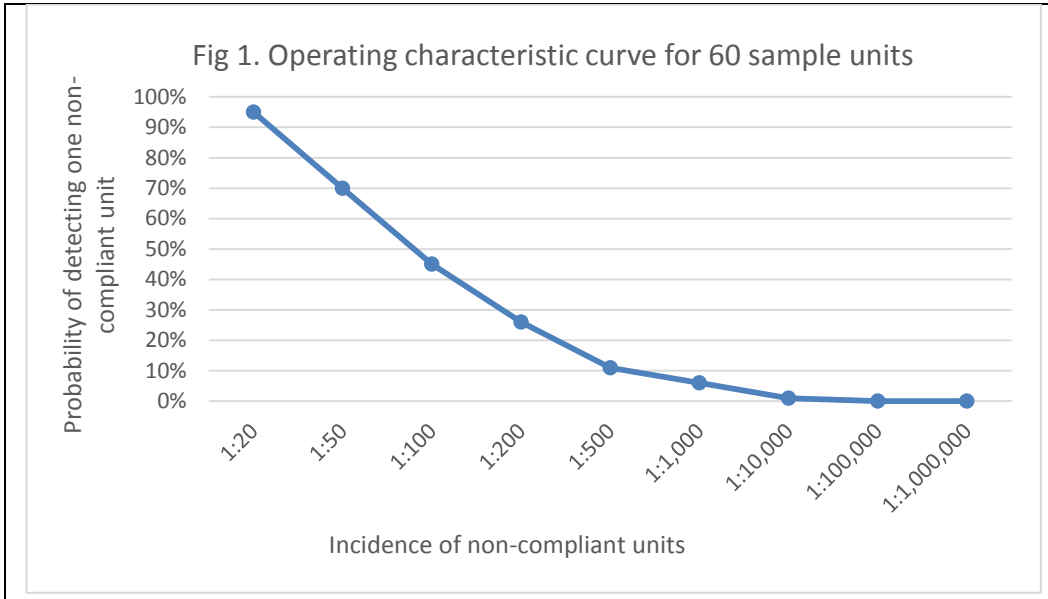
	Introduction The safety and suitability of foods is achieved through production control systems, such as HACCP. Sampling and testing of trade lots can provide added assurance that control systems are effective, however, such sampling has limitations and is not an effective control by itself. Government, industry, and Codex standards list performance criteria that are based on food safety and suitability objectives and appropriate levels of protection[2]. Performance criteria listed in food standards should be readily achievable using recognized industry-specific GMP and HACCP systems. Food safety and suitability performance criteria (e.g. chemical, biological and physical hazards, nutritional content, filth, decomposition) typically range from one non-conforming unit per thousand units, to one per million units. Food quality performance criteria (e.g. color, texture, size, workmanship defects) typically range from 1 per 100 units up to "six sigma" (3.4 defects per million) for industry production objectives. The incidence of non-compliant units expected from production control systems is much lower than can be verified using practical sampling plans because the number of samples that can be tested is limited by the costs of sampling, testing, and product destruction. Therefore, sampling does not ensure performance criteria are met; however, it can detect major food safety and quality issues affecting a large proportion of a lot. Attribute sampling plans Two-class attribute sampling plans are commonly used by customers to help verify the effectiveness of supplier control systems, and are used to accept or reject individual lots. Two-class attribute plans are based on binomial probability (an individual sample unit is either compliant or non-compliant). Attribute plans have the advantage of being distribution-free (e.g. they do not assume the level of a safety or suitability characteristic follows a specified parametric distribution such as the normal distribution with a known standard deviation). Table 1 shows the performance of zero-acceptance number, two-class attribute sampling plans for different sample sizes. With zero-acceptance number plans, detection of one or more non-compliant sample units indicates the lot is unacceptable. In Table 1, the number of sample units per lot is limited to 60 because it is generally impractical to sample and test more than 60 samples. The weakest performance criteria listed in Table 1 is 1:20 because criteria allowing more than one non-compliant unit per 20 units would indicate a defect of limited concern.
USA Table 1. Probability of detecting one or more non-compliant units (as a percentage) based on the incidence of non-compliant units in the lot and the number of sample units tested.	

[2] CAC Procedural Manual, 25th edition

Number of units tested	Incidence of non-compliant units (performance criteria)								
	1:20 (5%)	1:50 (2%)	1:100 (1%)	1:200 (0.5%)	1:500 (0.2%)	1:1,000 (0.1%)	1:10,000 (0.01%)	1:100,000 (0.001%)	1:1,000,000 (0.0001%)
3	14%	6%	3%	1%	1%	0%	0%	0%	0%
6	26%	11%	6%	3%	1%	1%	0%	0%	0%
10	40%	18%	10%	5%	2%	1%	0%	0%	0%
15	54%	26%	14%	7%	3%	1%	0%	0%	0%
20	64%	33%	18%	10%	4%	2%	0%	0%	0%
25	72%	40%	22%	12%	5%	2%	0%	0%	0%
30	79%	45%	26%	14%	6%	3%	0%	0%	0%
40	87%	55%	33%	18%	8%	4%	0%	0%	0%
50	92%	64%	39%	22%	10%	5%	0%	0%	0%
60	95%	70%	45%	26%	11%	6%	1%	0%	0%

Table 1 shows that examining up to 60 sample units has little to no probability of detecting an unacceptable lot for typical food safety and suitability performance criteria (1:1,000 to 1:1,000,000). Zero-acceptance number attribute plans are the most efficient, however, a practical plan can only reliably detect (with 95% confidence) control failures resulting in an incidence of 1 in 20 non-compliant units.

An ‘operating characteristic curve’ (‘OC curve’) can be created for a given sampling plan (single row in Table 1). Figure 1 shows an OC curve for 60 sample units, plotting the probability of detection (lot rejection) against the ‘incidence of non-compliant units’.



'Consumer's risk', and 'producer's risk', are traditional terms used in the sampling literature for the probability of accepting an unacceptable lot, and the probability of rejecting an acceptable lot, respectively. When applying performance criteria to sampling plans, the 'consumer's risk point' is the incidence of non-compliant units in a lot that the ideal sampling plan should detect with high probability. For example, to reliably detect (95% probability) lots that have 1:20 non-compliant units, requires testing at least 60 sample units (see Table 1 and Figure 1). For this sampling plan (i.e. 60 samples) the 'consumer's risk point' falls on the OC curve at 'incidence' = 1:20, and 'probability of detection' = 95%.

High probability (or confidence) is generally considered to be 95% or better. For such a confidence level, an OC curve will not pass through the 'consumer's risk point' for performance criteria more stringent than 1:20, when limited to 60 samples.

The 'producer's risk point' is the incidence of non-compliant units that should be accepted with high probability (e.g. 95%) based on the practical limitations of the industry's control systems. Table 1 shows that the sampling plan that detects 1:20 non-compliant units with 95% probability (using 60 samples), simultaneously has fairly low probability (6%) of rejecting lots with 1:1,000 non-compliant units, or inversely, fairly high probability (94%) of accepting lots with 1:1,000 non-compliant units. Therefore, the producer should maintain the level of non-compliant units below 1:1,000 to have high probability (>94%) of lot acceptance, when a lot is sampled and tested.

If the 60-sample, zero-acceptance number plan's default 'producer's risk point' ('incidence' = 1:1,000, and 'probability of detection' = 6%) is considered too stringent for producers to practically achieve, then the sampling plan can be adjusted (OC curve steepened) to include a more lenient 'producer's risk point' by allowing one or more non-compliant sample units to be detected before rejecting a lot (acceptance number > 0). However, increasing the acceptance number above zero requires a large increase in the total number of samples to maintain the same 'consumer's risk point'. If the acceptance number is increased from zero to one, in this example, then 93 instead of 60 samples are required to continue to reliably detect lots with 1:20 non-conforming units (with 95% probability).

If the characteristic being measured is unrelated to consumer health, then a compromise sampling plan requiring fewer samples may be considered, such as using 80% probability of detection (instead of 95%) for the 'consumer's risk point' and 20% probability of detection (instead of 5%) for the 'producer's risk point'.

When 60, or less, samples are used to detect a health hazard (performance criteria 1:1,000 to 1:1,000,000) the probability of detecting unacceptable lots is already low (< 6%) and the 'producer's risk point' is not a consideration, unless food security (starvation/malnutrition) is a balancing concern.

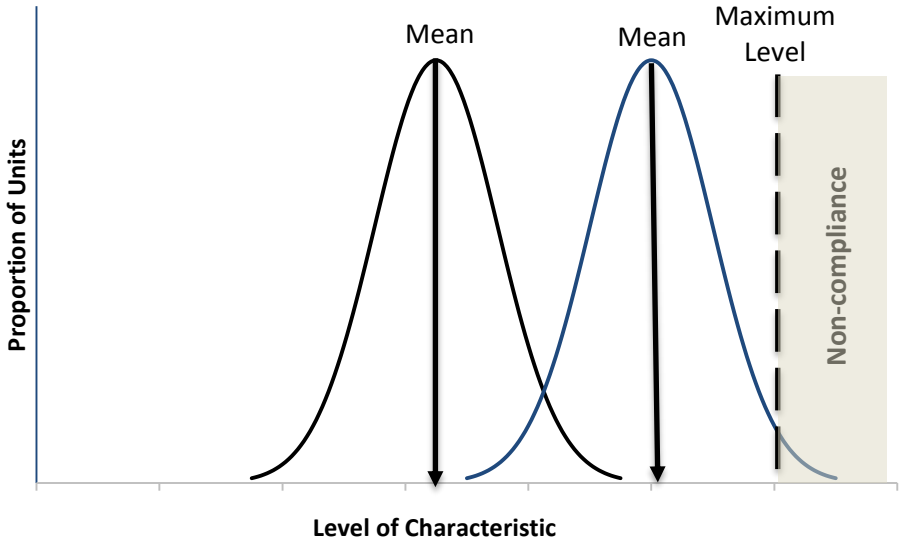
Variables Sampling Plans

'Variables sampling plans' are sometimes used in situations where the safety or quality characteristic measured is quantitative and follows a known distribution and a known standard deviation. Known distributions and standard deviations are most likely to occur when regularly sampling finished product within a production facility, or in an analogous situation.

Normal distribution with known standard deviation

Most variables sampling plans assume the measured variable (or log10 of the variable) follows a normal distribution. When the standard deviation of the variable's distribution is known and does not change from lot to lot, the incidence of non-compliant units in a lot can be calculated based on the average level of the tested safety or quality characteristic in the lot examined (see Figure 2).

Figure 2. Two normally distributed populations with different means and the same standard deviation.



The mean values shown in Figure 2 represent average levels in lots. Since the standard deviation (width of the curves) is presumed constant, the lot average determines the percentage of product that will be over the maximum level. When the lot average indicates the lot does not meet performance criteria, the lot is rejected. Note that variables sampling plans are designed to reject lots at average levels that are within acceptable levels for individual units, and none of the sample units tested are required to exceed the maximum (or minimum) acceptable level in order to reject a lot.

The average level can be calculated mathematically, or, depending on the type of hazard (e.g. chemical), it can be attained mechanically by homogenizing the sample units in a blender before testing (aggregate sample). Aggregate samples save costs by testing one sample instead of each sample unit separately; however, a single aggregate sample does not provide information on variability to help confirm the presumed known standard deviation.

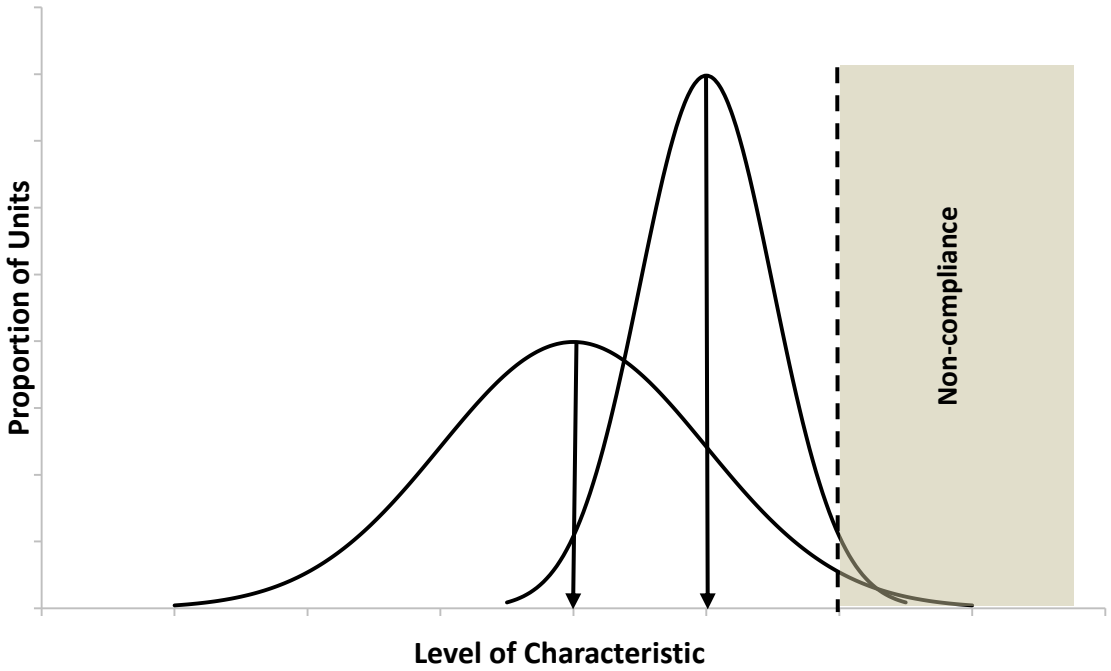
When distribution presumptions are met, variables sampling plans require fewer samples than attribute sampling plans for the same probability of detecting unacceptable lots. This is because variables sampling plans use the exact level of the variable in each sample unit and a fixed distribution shape, while attribute sampling plans use an assessment of whether the variable is over or under a limit and do not require a fixed distribution shape.

Variables sampling plans are not used for typical trade lots because the distribution shape and standard deviation are likely to change from lot to lot. Trade lots are received from various suppliers, and even when purchased from the same supplier, shipping containers may contain mixtures of items produced over multiple days from different raw materials. Standard deviations for product hazards, such as histamine in tuna, vary widely from lot to lot. Generally, the higher the incidence of non-compliant units in a lot, the larger the standard deviation. Variables sampling plans based on a fixed standard deviation are not applicable if the standard deviation changes because the acceptable lot average is proportional to the standard deviation. Attribute sampling plans are appropriately recommended for trade lot applications because they are not affected by the distribution shape or the standard deviation.

Normal distribution with unknown standard deviation

When the standard deviation is unknown there is a type of variables sampling plan that can be used where the standard deviation is estimated for each lot separately. The lot standard deviation is estimated from the individual sample units tested from the lot. When the lot standard deviation is smaller (the narrower distribution in Figure 3), the acceptable lot average is closer to the assigned maximum level; and when the lot standard deviation is larger (the wider distribution in Figure 3), the acceptable lot average is farther from the maximum level. Because the acceptable lot average changes with each lot tested, this approach is difficult to use for trade standards. However, it can be used to screen lots for potential problems.

Figure 3. Two normally distributed lots with different means and different standard deviations. Both lots have the same incidence of non-compliant units indicated by the area to the right of the dashed line, however, they have different acceptable averages.


Proportion of Units
Non-compliance
Level of Characteristic