STATEMENT ON ASTM E2548-16

To: OSAC Program Office
From: David Kaye
Subject: Statistical Terminology in ASTM E2578-16
Date: November 6, 2020

Introduction

In 2016, six OSAC Legal Resource Committee members submitted public comments opposing the addition of ASTM E2548-11e1 (Standard Guide for Sampling Seized Drugs for Qualitative and Quantitative Analysis) to the OSAC Registry. Despite these comments and criticism from other OSAC units, OSAC added the standard to the Registry (effective April 3, 2017).

ASTM International made minor changes to the 2011 Standard Guide in ASTM E2548-16. This superseding version has been proposed for addition to the Registry. It is difficult to say whether the 2016 version should be added to the Registry. On the one hand, it contains several improvements, and it can be argued that an improved version—no matter how slight the improvement—should replace the 2011 version on the Registry. On the other hand, a negative outcome would underscore the need to finally arrive at a standard that adequately addresses the concerns that lawyers, statisticians, and psychologists have voiced.[1]

Without trying to resolve the Registry-approval question, these comments reiterate several of the technical objections to the standard and offer further observations so that they will remain clearly in sight. The major, interrelated, problems in ASTM E2548-16 are the failure: (1) to state clearly that probability sampling is *required* to permit valid inferences about population features (the technical term is "parameters"); (2) to define basic terms; (3) to use standard terms (in the statistics literature) for concepts; (4) to give some guidance on what values for the confidence coefficient or related criteria are scientifically acceptable in picking sample sizes using standard statistical methods; and (5) to offer guidance on how to express sampling uncertainty in either the frequentist and Bayesian frameworks to which the Standard Guide alludes.

Because the Standard Guide's discussion of statistical reasoning is vague, laboratories will have to consult other sources to know what to do, making the Guide useful primarily as a bibliography. OSAC and ASTM should work toward having an easily understood document for drug chemists who want to develop sampling plans and present the results of tests on a sample of items correctly. Therefore, the best strategy might be to bring in a statistical expert on sampling theory and methods to reorganize and rewrite this Standard Guide rather than to continue making patchwork changes and revisions. Nevertheless, the remainder of this statement contains notes on some of the specific wording in ASTM E2548-16 that is problematic. Some of the remarks are editorial and minor. Others are editorial and major. Some are conceptual and substantive.

[1] A paragraph at the end of the Guide states that it "must be reviewed every five years and if not revised, either reapproved or withdrawn." The five-year review should occur next year, making it unclear why it is urgent to act on a 2016 version whose withdrawal or replacement is imminent.

Detailed Remarks

1.1 This guide covers minimum considerations for sampling of seized drugs for qualitative and quantitative analysis.

← This is not a precise statement of the scope of the Standard Guide. What is a "minimum consideration"? Depending on how one construes the phrase "minimum consideration for sampling," the statement of scope is not complete. The Guide also has sections on the documentation (§ 7) and reporting (§ 8) of the sampling part of the laboratory's work. These exceed the stated scope. At the same time, the Guide lacks a parallel section on testimony on the nature and outcomes of the laboratory's sampling. The scope section should outline the full scope and the important topics that are left out.

3.2 The principal purpose of sampling in the context of this guide is to answer relevant questions about a population by examination of a portion of the population. For example: What is the net weight of the population? What portion of the units of a population can be said to contain a given drug at a given level of confidence?

← The first sentence is not quite correct. The purpose of *sampling* is merely to produce a representative sample for analysis. The combination of sampling, analysis (measurement), and interpretation gives an answer to (or input that helps to answer) a legally relevant question without having to analyze every unit in the population.

"Population" should be defined in a new section that defines technical terms. The definition can be taken from one of the references or a statistics text. Same for "unit."

The question "What is the net weight of the population?" should be "What is the net weight of a proscribed substance in the population?".

"Level of confidence" should be defined. Does it refer only to the value of a confidence coefficient in a frequentist interval estimate? To a posterior probability in a Bayesian credible region? Both?

3.3 By developing a sampling strategy and implementing appropriate sampling schemes, as illustrated in Fig. 1, a laboratory will minimize the total number of required analytical determinations, while ensuring that all relevant legal and scientific requirements are met.

← "Sampling strategy" and "sampling scheme" should be defined. Neither phrase is common in statistics texts. There may be better terms.

This subsection presumes that there is no conflict between scientific and legal requirements. In any case, if "appropriate" means a procedure that reduces the number of units sampled to the minimum that is reasonably likely to estimate a parameter with scientifically and legally satisfactory precision, then the sentence should be rewritten to express that objective.

4.1 An appropriate sampling strategy is highly dependent on the purpose of the investigation, a customer's request, and the anticipated use of the results. Laws and legal practices form the foundation of most strategies and shall be taken into account when designing a sampling scheme. Therefore, specific sampling strategies are not defined in this guide.

← The paragraph concludes with a non sequitur. Specific sample designs can and should be defined and discussed in a Standard Guide. Various legal constraints and customer requests can be enumerated, and guidelines can be provided for the enumerated situations. Doing so would supply more useful guidance than self-evident remarks such as "legal practices … shall be taken into account."

4.2.1 Sampling may be statistical or non-statistical. NOTE 1—For the purpose of this guide, the use of the term statistical is meant to include the notion of an approach that is probability-based.

← Normally, the word "statistical" means "relating to the use of statistics," and a statistic is a number derived from or characterizing data (such as a batting average). In general (and as in the example of a batting average), the data need not be acquired by a method that involves the application of the mathematical theory of probability at some stage (if that is what "an approach that is probability based" means). Of course, one can choose to redefine a term to depart from its normal usage in science and ordinary life, but there is no reason to do so when a perfectly good term (in this case, "probability sampling") is available in the statistical literature.

4.2.1.1 In many cases, a non-statistical approach may suffice. The sampling plan shall provide an adequate basis for answering questions of applicable law. For example, "Is there a drug present in the population?" "Are statutory enhancement levels satisfied by the analysis of a specified number of units?"

← The term "sampling plan" should be defined. How does it differ from a "sampling scheme," "sampling strategy," and "sampling procedure"? Figure 1 has all these terms in it, but most of them are not defined explicitly and clearly. ("Sampling scheme" is defined in § 5.1.)

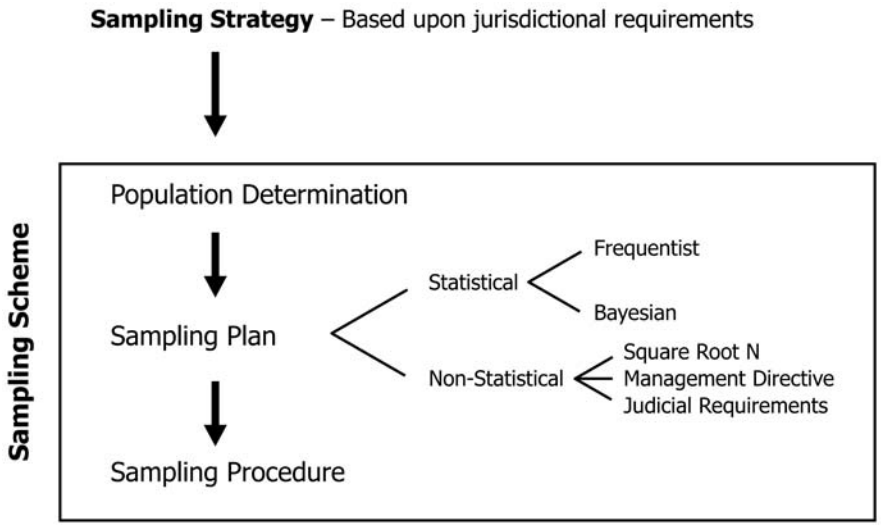
4.2.1.2 If an inference about the whole population is to be drawn from a sample, then the plan shall be either statistically based or have an appropriate statistical analysis completed and limits of the inference shall be documented.

← The phrase "or have an appropriate statistical analysis completed" has been added since 2011. But what "appropriate statistical analysis" will permit a statistical inference from a sample to the population in the absence of a probability sample? In some research, non-probability samples are treated as if they were probability samples, but it is not necessary to make that analogy if probability sampling is feasible, as it seems to be here. (If it is not always feasible, those situations should be identified.)

4.2.2 Selected units shall be analyzed to meet Practice E2329 if statistical inferences are to be made about the whole population.

← ASTM E2329 describes minimum requirements for qualitative analysis. It is not obvious why these requirements for identifying compounds do not apply to the analysis of every unit in a census of the population. Moreover, this provision seems beyond the scope of a standard on sampling. It concerns measurements or observations on the items in a sample.

FIG. 1. Relationship of Various Levels Required in Sampling



← Previously, the leaves in the top “statistical” branch were "Hypergeometric," "Bayesian," and "Other probability-based approaches." These have been trimmed to “Bayesian” and “Frequentist.” This is an improvement, since “Hypergeometric” is merely the name of a probability distribution, and not a form of sampling. (When sampling without replacement, the probability that a number *X* of units with a discrete feature will be picked in *N* draws follows the hypergeometric distribution. This is true for Bayesians and frequentists alike.)

The problem that remains in this branch is that neither “Bayesian” nor “frequentist” describes a sampling plan. For example, the results for a simple random sample (SRS) can be analyzed from either perspective, and the sample size can be pre-established according to computations with either methodology (or via likelihood computations, for that matter). Presumably, a sampling procedure that implements the SRS design would be a “sampling plan.” But the SRS design itself is neither Bayesian nor frequentist. In addition, there is a third school of statistical thought, likelihood theory, that could be used in analyzing sample data.

The leaves of the “Non-Statistical” branch are “Square Root N,” “Management Directive,” and “Judicial Requirements.” The first leaf refers to a procedure for establishing the size of a sample. The square root of *N* assuredly is a statistic, and unless the population size exceeds 9, the method is to draw an SRS (see United Nations Office on Drugs and Crime, Recommended Methods for Testing Opium, Morphine and Heroin: Manual for Use by National Drug Testing Laboratories, p.21). Consequently, the “square root method” can be used to obtain a sample statistic that is a statistically sound estimator of a population parameter.

The next two leaves can apply to any choice of a sampling design that comes from a manager or a court. The design could be statistically optimal or wasteful, but its statistical properties do not make it “non-statistical.” Perhaps the branch is supposed to be “not suitable for statistical inference,” but that does not lead to the three leaves either.

In sum, Figure 1 muddies the waters. It should be abandoned or redrawn after more appropriate and better defined terminology is in place.

5.2.1 The population determination shall take into account all typical forms and quantities in which exhibits may appear.

What is an “exhibit”? The term appears nowhere else in the Guide. Why should atypical forms and quantities (of what?) not be considered along with typical ones?

5.2.2 A population can consist of a single unit or multiple units.
5.2.3 A multiple unit population shall consist of items that are similar in relevant visual characteristics (for example, size, color, shape, etc.).

← The wording suggests that material packaged in containers of different sizes or color is not a population of legal interest. In this form, the material is neither a “single unit” nor a “multiple unit” population. The statistical question may concern a feature of all the material, and that collection would be the population. It appears that in this standard, a “population” is a sampling frame that is constructed from the population. The literature on attribute and variable sampling as well as multistage sampling should be consulted. In sampling theory, a “sampling unit” is not necessarily a single element of the population. There are smaller units of elements at each stage.

5.4 When a single unit or bulk population is to be analyzed, the issue of homogeneity shall be addressed within the sampling plan.

← Now it appears that there is a third kind of population—a “bulk population” that is neither a single unit nor a combination of units. The definitions of all these populations should be in a definitions section.

5.4.1 One sample is sufficient if the bulk material is homogeneous. Analysts can make bulk material homogeneous.

← Normally, a sample is the subset of elements drawn from the sampling frame that represents the population of interest. Of course, “sample” has another meaning—a single chunk of material for instrumental or visual analysis—but that is not necessarily a “sample” as the term is used in sampling theory. The sample from the population can consist of many items. Substituting “specimen” for “sample” in this section would avoid possible confusion.

5.5 For a multiple unit population, the sampling plan may be statistical or non-statistical.

← See supra comment on Figure 1.

5.5.1 Statistical approaches are applicable when inferences are made about the whole population. For example:
The probability that a given percentage of the population contains the drug of interest or is positive for a given characteristic.
The total net weight of the population is to be extrapolated from the weight of a sample.

← Normally one would speak of the population percentage as a parameter whose value is being estimated from sample data. The estimate is an inference. A hypothesis about the true value would be another inferential statement. The “probability that a given percentage of the population contains the drug of interest or is positive for a given characteristic,” if based on the sample data and a prior distribution, would be called a posterior probability for the inference. The posterior probability cannot

be computed with frequentist methods for inference, and this fact should be made clear in the Standard Guide (and included in reports that contain coverage probabilities).

The second example is not the inference; it is the objective of the study. An example of an inference would be "Extrapolating from the sample data, the total net weight of a controlled substance in the population is 500 grams."

5.5.1.1 Published examples of statistical approaches involving general considerations:
(1) Practice E105.
(2) Practice E141.
(3) Terminology E1732.

← A glossary or dictionary of terms is not an example of an approach.

(4) Guidelines on Representative Drug Sampling.[4]
[4] 2009 UNODC/ENFSI Guidelines on Representative Drug Sampling, European Network of Forensic Science Institutes (ENFSI), 2004, http://www.ENFSI.org.

← A 2009 document cannot be published in 2004.

(5) ISO 3534-1.
(6) ISO 3534-2.

← A glossary or dictionary of terms is not an example of an approach, but maybe some of the definitions contain such examples.

5.5.2 Non-statistical approaches are appropriate if no inference is to be made about the whole population.
5.5.2.1 Published examples:
(1) The "square root method."[7]
(2) Methods listed in "Arbitrary Sampling" in *Guidelines on Representative Drug Sampling*.[4]

← The "square root method" of (1) is a form of "arbitrary sampling" according to the *Guidelines* cited in (2). Therefore, having the two categories of "non-statistical approaches" makes little sense (even if one can construe using an SRS with the various "square root methods" listed in the EU document for choosing the sample size as "non-statistical").

5.5.2.3 A non-statistical sampling approach may allow an inference about the population. If a single population has been randomly sampled, the data may allow an inference to be drawn by (1) determining and reporting a confidence interval for an inferred population parameter (for example: weight or tablet count); or by (2) retrospectively using the results in a statistical model and determining the resulting probabilities and level of confidence.

← This paragraph is a real puzzler. The essence of probability sampling is that the probability that each and every element in the sampling frame is known. Knowing these probabilities allows one to use a statistical model to determine the sampling distribution of the estimator. That, in turn, leads to a standard error, a point estimate, and thus a confidence interval. If "non-statistical sampling" does not use probability sampling, how is one supposed to determine the standard error? The answer might be to analogize the approach actually used to a probability sampling method of some sort. But that interpretation of the paragraph is contradicted by "(2) retrospectively using the results in a statistical

model and determining the resulting probabilities and level of confidence." (1) is already retrospective and uses a statistical model to produce the confidence interval, so what is (2) supposed to add? Furthermore, (2) is obscure at best. A "level of confidence" is a coverage probability. As such, what are "the resulting probabilities"?

5.6.1 Establish the procedure for selecting the number of units that will comprise a sample.

← This sounds prescriptive, but the relative merits of different procedures have, strangely, not been discussed in this Standard Guide. Accordingly, the document supplies no guidance beyond a general admonition to read the literature mentioned in it and to do the right thing. What would that be? The UN report, for example, contends that even though the square-root sample-size determination lacks a sound theoretical foundation, it gives reliable results. The EU report, which describes this rule-of-thumb as "arbitrary," acknowledges that it does "work well in many situations," and explicates a number of variations on it. But statisticians have been known to dismiss it out of hand. E.g., Alan Julian Izenman, Statistical and Legal Aspects of the Forensic Study of Illicit Drugs, 16 Stat. Sci. 35–57, 47 (2001) ("The popularity of the square root rule, despite the lack of theoretical support for this rule, shows how an unfounded rule-of-thumb can be established in the practice of a particular field."). What guidance does the Standard Guide have to offer?

5.6.1.1 For non-statistical approaches, select a sample appropriate for the analytical objectives.

← There is no indication in the Standard Guide of how to go about doing this and no clear statement of what "analytical objectives" are. Usually, "analysis" in the Standard Guide refers to laboratory analysis (particularly § 6), but in this subsection the phrase "analytical objectives" seems to be a veiled reference to doing what management or the law dictates. Or, if that is not what the words point to, does not the same advice apply to "statistical approaches"? Surely, in all situations, one should select "a sample appropriate for the analytical objectives." The reader is left uninformed by the Guide.

5.6.1.2 For statistical approaches, random sampling shall be conducted.

← Given that probability sampling always is possible in this context (at least approximately), why would one ever want anything less? For example, if the sample size n is determined by the formula $n = N/10$, where N is the population size (based on an opinion in which some judge at some time said that he or she thought that samples of less than 10% could not be trusted), are we in the 5.6.1.1 realm of "non-statistical approaches"? The selection method should use a randomizing mechanism regardless of a "non-statistical" approach to determining the sample size.

5.6.2.1 A random sample is one selected without bias and where each item has an equal chance of being selected.

← The words "without bias" seem like they represent a separate requirement, but they are merely superfluous. The reason to give an item in a sampling unit (which could be the sampling frame) an equal probability of being selected is to avoid bias (and give estimates of sampling error a solid foundation). The sentence could be broken into the three different thoughts: "A simple random sample is one in which each item in the sampling frame has an equal probability of being selected. Using a randomizing procedure to equalize the probability avoids selection bias and supplies the foundation for estimating sampling error."

Why is simple random sampling the only acceptable form of probability sampling?

5.6.2.1 (cont'd) Computer generated random numbers or random number tables are commonly employed for such tasks and these should be included in the sampling plan.

5.6.2.2 Random sampling of items using random number tables may not be practical in all cases. In these instances, an alternate sampling plan shall be designed and documented to approach random selection. A practical solution involves a "black box" method, which refers to one that will prevent the sampler from consciously selecting a specific item from the population (that is, all units are placed in a box and the samples for testing are selected without bias). …

← 5.6.2.1 requires a sampling plan that provides for selection according to random numbers even when their use is impractical. Then 5.6.2.2 says not to use that part of the plan in some cases. The final clause of 5.6.2.1 should be "and these should be part of the sampling plan unless a statement showing that they are impractical is included in the written plan."

Why are "random number tables" but not pseudo-random number generators the subject of 5.6.2.2?

5.7 Sample Reduction—Sample reduction may be applied in cases where the weight or volume of the selected units is too large for laboratory analysis (Fig. 2, insert A).

← "Sample reduction" should be defined. Figure 2 does it to some extent (by listing the names of three methods for reducing the quantity), but a section with a general definition would improve readability. Again, in the figure, "sample" is used in the laboratory-analysis sense rather than the sampling-from-a-population sense. A word or phrase like "specimen" or "analytical unit" would be better for the former kind of sample.

6. Analysis
6.1 Statistically Selected Sample(s)—In accordance with 4.2.2, it is recommended that each unit comprising the sample be analyzed to meet Practice E2329 if statistical inferences are to be made about the whole population.
6.2 Non-statistically Selected Sample(s)—Practice E2329 shall be applied to at least one unit of the sample.

← See comment on § 4.2.2.

7. Documentation
7.1 Inferences drawn from the application of the sampling plan and subsequent analyses shall be documented.

← This section is exceedingly skimpy. It does not even cover the documentation mentioned in § 4.2.1.2 and § 5.6.2.2

8. Reporting
8.1 Sampling information shall be included in reports.

What "sampling information" should be reported? The subsections that follow (8.1.1 and 8.1.2) refer to nothing more than a statement of the results. They do not discuss reporting the parts of the sample design and selection mechanisms that led to the testing.

8.1.1 Statistically Selected Sample(s)—Reporting statistical inferences for a population is acceptable when testing is performed on the statistically selected units as stated in 6.1 above. The language in the report must make it clear to the reader that the results are based on a probability-based sampling plan.
8.1.2 Non-Statistically Selected Sample(s)—The language in the report must make it clear to the reader that the results apply to only the tested units. For example, 2 of 100 bags were analyzed and found to contain Cocaine.

← The Guide is confusing in its description of what qualifies as "a probability-based sampling plan." For example, it appears from Figure 1 and parts of the text noted above that if the sample size is not chosen by using probability theory at that stage, the result is a "non-statistically selected sample." The report therefore could not make any extrapolation to the population from an SRS with an ad hoc sample size. But using an unnecessarily large sample—which is what using a statistically optimal sample size guards against—does not undermine the validity of a statistical inference from the SRS. The laboratory should be able to make inferences from probability samples regardless of the method for choosing the sample size.

When the sample is not a probability sample—one drawn so as to give every element in the sampling frame a known probability of selection—an ampliative inference is of questionable validity, and the demand that the report "make it clear to the reader that the results apply to only the tested units" is reasonable. But merely stating that *x* out of *N* bags "were analyzed and found to contain Cocaine" does not make it clear that the sample proportion *x*/*N* might deviate greatly from the population proportion and that, without simple random sampling, statistical theory cannot be used to estimate the sampling error. Moreover, although the "*x* of *N*" format for reporting is just as appropriate for estimates from probability samples, § 8.1.1 does not mention it. It only addresses the need to report that "a probability-based sampling plan" was used. It overlooks the issue of how to report the results themselves.

9. Keywords
9.1 analytical method; qualitative method; quantitative method; sampling; seized drug analytical method

← The only keyword that has to do with sampling is "sampling." For a Standard Guide on sampling, that is shocking.

OSAC Seized Drugs Subcommittee Response to Public Comment on the Replacement of ASTM E2548-11e1 with E2548-16 on the Registry:

ASTM E2548-11e1 "Standard Guide For Sampling Seized Drugs for Qualitative and Quantitative Analysis" is currently on the OSAC Registry. This standard has been superseded at ASTM by E2548-16; as such, the OSAC Seized Drugs Subcommittee submits ASTM E2548-16 to replace the historical version of the standard on the OSAC Registry. Before discussing the merits of E2548-16, we would like to make clear that we agree with the general sentiment expressed in the public comments that this standard would be well-served by a comprehensive revision in collaboration with statisticians who have specific expertise in sampling and attorneys who have interest and experience with these issues, such as David Kaye (the LTG chair). Indeed, to that end we have opened a new item in ASTM (WK75229) to begin the redrafting process.

We feel it is important to replace the superseded E2548-11e1 document with the E2548-16 version on the Registry at this time because several improvements were made in the E2548-16 document and it is in the best interest of the forensic community for OSAC to refer individuals to the current version of the document. The improvements made in E2548-16 include:

- Random sampling is necessary when making an inference about a population. This is now a requirement in E2548-16, but was only a recommendation in the E2548-11 version.
- E2548-11 does not provide information on the possibility of making an inference to the population from non-statistical sampling whereas E2548-16 does.
- Figure 1 has been improved in E2548-16, as acknowledged by Commenter 1.

The challenge we face in terms of timing is that the process of revising the standard through both the OSAC and ASTM processes will take 3-5 years and we feel the addition of the compulsory language regarding random sampling and the improvement to figure 1 are important enough changes to support replacement of E2548-11e1 with E2548-16 on the Registry.

Public comments received regarding this action have been transcribed below along with the subcommittee adjudication and response to each item.

Adjudication of Public Comments:

<u>Commenter 1:</u>

1. Neither ASTM E2548-11e1 nor its successor, ASTM E2548-16, belong on "a repository of high-quality, technically sound published ... standards for forensic science." The basis for this conclusion can be found in the attached memorandum. If the choice is between allowing ASTM E2548-11e1 to remain on the Registry and replacing it with ASTM E2548-16 pending a major overhaul, however, the pragmatic question of whether the 2016 version belongs on the Registry is more complex.

   Subcommittee Response: We understand this comment but believe the improvements made in E2548-16 over the superseded E2548-11e1 version, in addition to the benefit to the community for

OSAC to refer laboratories to the current version of a standard, warrant the replacement of the historical version of the standard with the revised version of the standard on the Registry.

2. In 2016, six OSAC Legal Resource Committee members submitted public comments opposing the addition of ASTM E2548-11e1 (Standard Guide for Sampling Seized Drugs for Qualitative and Quantitative Analysis) to the OSAC Registry. Despite these comments and criticism from other OSAC units, OSAC added the standard to the Registry (effective April 3, 2017). ASTM International made minor changes to the 2011 Standard Guide in ASTM E2548-16. This superseding version has been proposed for addition to the Registry. It is difficult to say whether the 2016 version should be added to the Registry. On the one hand, it contains several improvements, and it can be argued that an improved version—no matter how slight the improvement—should replace the 2011 version on the Registry. On the other hand, a negative outcome would underscore the need to finally arrive at a standard that adequately addresses the concerns that lawyers, statisticians, and psychologists have voiced.[1]

   Subcommittee Response: Unfortunately, the revisions to ASTM E2548 were already underway at ASTM when E2548-11e1 was submitted for registry approval in 2016, and the comments received through the OSAC process were not able to be incorporated at the SDO at that time. A work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group and invite a statistician and an LRC member into the task group to participate in the drafting of the revised document. The challenge faced is that the process of revising the standard through both the OSAC and ASTM processes will take 3+ years and we feel it is important to replace the superseded E2548-11e1 document with the E2548-16 version on the Registry since several improvements were made in the E2548-16 document and it is in the best interest of the forensic community for OSAC to refer individuals to the current version of the document.

3. Without trying to resolve the Registry-approval question, these comments reiterate several of the technical objections to the standard and offer further observations so that they will remain clearly in sight. The major, interrelated, problems in ASTM E2548-16 are the failure: (1) to state clearly that probability sampling is *required* to permit valid inferences about population features (the technical term is "parameters"); (2) to define basic terms; (3) to use standard terms (in the statistics literature) for concepts; (4) to give some guidance on what values for the confidence coefficient or related criteria are scientifically acceptable in picking sample sizes using standard statistical methods; and (5) to offer guidance on how to express sampling uncertainty in either the frequentist and Bayesian frameworks to which the Standard Guide alludes.

   Subcommittee Response: Persuasive. However, one of the major improvements to E2548-16 is that 5.6.1.2 now requires random sampling for statistical sampling approaches whereas this was only a recommendation in the E2548-11 version. The OSAC Seized Drugs Subcommittee supports a full revision of the standard. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group, and invite a statistician and an LRC member into the task group to participate in the drafting of the revised document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take

[1] A paragraph at the end of the Guide states that it "must be reviewed every five years and if not revised, either reapproved or withdrawn." The five-year review should occur next year, making it unclear why it is urgent to act on a 2016 version whose withdrawal or replacement is imminent.

3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

4. Because the Standard Guide's discussion of statistical reasoning is vague, laboratories will have to consult other sources to know what to do, making the Guide useful primarily as a bibliography. OSAC and ASTM should work toward having an easily understood document for drug chemists who want to develop sampling plans and present the results of tests on a sample of items correctly. Therefore, the best strategy might be to bring in a statistical expert on sampling theory and methods to reorganize and rewrite this Standard Guide rather than to continue making patchwork changes and revisions.

   Subcommittee Response: Not persuasive as to this version, but will be considered as part of the full revision of the standard. This particular ASTM standard is a guide. Per ASTM requirements, guides are a compendium of information or series of options that do not recommend a specific course of action. However, we do recognize the importance of sampling in forensic laboratories and the importance of providing easily understood guidance to the community. As such, the goal of the planned revisions to the standard will be to change the standard from a guide into a practice which per ASTM requirements is a set of instructions for performing one or more specific operations that does not produce a test result. As previously stated, we agree including an expert in sampling theory is the best course of action for drafting the new item.

5. "1.1 This guide covers minimum considerations for sampling of seized drugs for qualitative and quantitative analysis."

   This is not a precise statement of the scope of the Standard Guide. What is a "minimum consideration"? Depending on how one construes the phrase "minimum consideration for sampling," the statement of scope is not complete. The Guide also has sections on the documentation (§ 7) and reporting (§ 8) of the sampling part of the laboratory's work. These exceed the stated scope. At the same time, the Guide lacks a parallel section on testimony on the nature and outcomes of the laboratory's sampling. The scope section should outline the full scope and the important topics that are left out.

   Subcommittee Response: Editorial - does not critically impact the interpretation of the standard. The phrase "minimum considerations for sampling" is meant to encompass the entire sampling process to include the documentation and reporting of any sampling conducted by the forensic science practitioner. However, we do recognize the importance of testimony in relation to sampling and the planned revisions to the standard will incorporate a testimony section. The OSAC Seized Drugs Subcommittee has also begun work on a separate testimony standard.

6. "3.2 The principal purpose of sampling in the context of this guide is to answer relevant questions about a population by examination of a portion of the population. For example: What is the net weight of the population? What portion of the units of a population can be said to contain a given drug at a given level of confidence?"

   a. The first sentence is not quite correct. The purpose of *sampling* is merely to produce a representative sample for analysis. The combination of sampling, analysis (measurement), and interpretation gives an answer to (or input that helps to answer) a legally relevant question without having to analyze every unit in the population.

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

b. "Population" should be defined in a new section that defines technical terms. The definition can be taken from one of the references or a statistics text. Same for "unit."

Subcommittee Response: Not persuasive. Both terms are defined in ASTM E1732 "Standard Terminology Relating to Forensic Science", which is referenced in E2548-16. However, we do agree that a terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

c. The question "What is the net weight of the population?" should be "What is the net weight of a proscribed substance in the population?".

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard. Discussions with the commenter revealed the issue is with the "of the population" phrase and the suggested change is to "all items in the population".

d. "Level of confidence" should be defined. Does it refer only to the value of a confidence coefficient in a frequentist interval estimate? To a posterior probability in a Bayesian credible region? Both?

Subcommittee Response: Editorial. There is a discussion of the term in E1732, however, a terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

7. "3.3 By developing a sampling strategy and implementing appropriate sampling schemes, as illustrated in Fig. 1, a laboratory will minimize the total number of required analytical determinations, while ensuring that all relevant legal and scientific requirements are met."
   a. "Sampling strategy" and "sampling scheme" should be defined. Neither phrase is common in statistics texts. There may be better terms.

   Subcommittee Response: Editorial. These terms are used within seized drugs forensic laboratories, and although each term is discussed within the standard, we agree that a terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

   b. This subsection presumes that there is no conflict between scientific and legal requirements. In any case, if "appropriate" means a procedure that reduces the number of units sampled to the minimum that is reasonably likely to estimate a parameter with scientifically and legally satisfactory precision, then the sentence should be rewritten to express that objective.

   Subcommittee Response: Persuasive. In the seized drug discipline, not all sampling strategies involve estimating a parameter. We realize that the term 'sample' and 'sampling' as used in the standard may not be how the terms are defined in the statistics community. Seized-drug practitioners historically refer to both sample selection and sampling as "sampling strategies". Based on this comment, we believe the standard in its current form could be improved in the delineation of these two concepts. A terminology

section will aid in this understanding and this comment will be considered during the five-year revisions to E2548-16 (ASTM WK75229).

We recognize that terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

8. "4.1 An appropriate sampling strategy is highly dependent on the purpose of the investigation, a customer's request, and the anticipated use of the results. Laws and legal practices form the foundation of most strategies and shall be taken into account when designing a sampling scheme. Therefore, specific sampling strategies are not defined in this guide."

   The paragraph concludes with a non sequitur. Specific sample designs can and should be defined and discussed in a Standard Guide. Various legal constraints and customer requests can be enumerated, and guidelines can be provided for the enumerated situations. Doing so would supply more useful guidance than self-evident remarks such as "legal practices … shall be taken into account."

   Subcommittee Response: Not persuasive as to this version, but will be considered as part of the full revision of the standard. This particular ASTM standard is a standard guide. Per ASTM requirements, guides are a compendium of information or series of options that do not recommend a specific course of action. However, we do agree that the standard could be improved and provide additional guidance to the reader and the goal of the planned revisions to the standard will be to change the standard from a guide into a practice to provide additional information. Examples may also be considered for an appendix to the standard as well.

9. "4.2.1 Sampling may be statistical or non-statistical. NOTE 1—For the purpose of this guide, the use of the term statistical is meant to include the notion of an approach that is probability-based."

   Normally, the word "statistical" means "relating to the use of statistics," and a statistic is a number derived from or characterizing data (such as a batting average). In general (and as in the example of a batting average), the data need not be acquired by a method that involves the application of the mathematical theory of probability at some stage (if that is what "an approach that is probability based" means). Of course, one can choose to redefine a term to depart from its normal usage in science and ordinary life, but there is no reason to do so when a perfectly good term (in this case, "probability sampling") is available in the statistical literature.

   Subcommittee Response: Editorial. The term is defined in NOTE 1 as referring to probability-based sampling, however, we recognize that terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has

been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

10. "4.2.1.1 In many cases, a non-statistical approach may suffice. The sampling plan shall provide an adequate basis for answering questions of applicable law. For example, "Is there a drug present in the population?" "Are statutory enhancement levels satisfied by the analysis of a specified number of units?"

    The term "sampling plan" should be defined. How does it differ from a "sampling scheme," "sampling strategy," and "sampling procedure"? Figure 1 has all these terms in it, but most of them are not defined explicitly and clearly. ("Sampling scheme" is defined in § 5.1.)

    Subcommittee Response: Editorial. These terms are used within seized drugs forensic laboratories, and although each term is discussed within the standard, a terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

11. "4.2.1.2 If an inference about the whole population is to be drawn from a sample, then the plan shall be either statistically based or have an appropriate statistical analysis completed and limits of the inference shall be documented."

    The phrase "or have an appropriate statistical analysis completed" has been added since 2011. But what "appropriate statistical analysis" will permit a statistical inference from a sample to the population in the absence of a probability sample? In some research, non-probability samples are treated as if they were probability samples, but it is not necessary to make that analogy if probability sampling is feasible, as it seems to be here. (If it is not always feasible, those situations should be identified.)

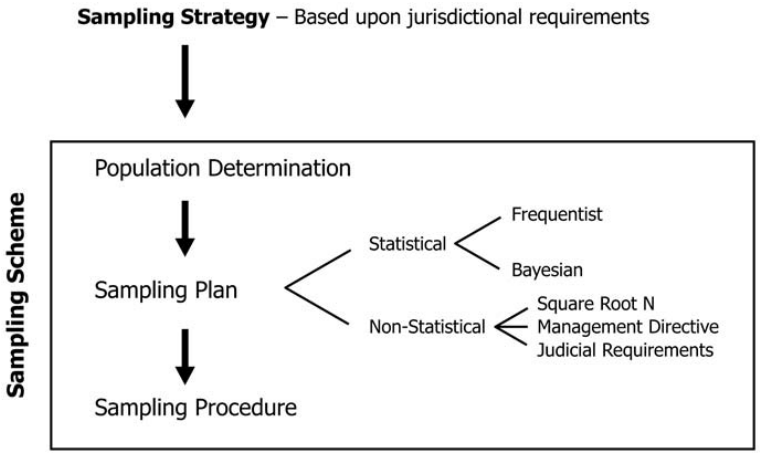
    Subcommittee Response: Persuasive. We agree that the standard could be improved and provide additional guidance to the reader and the goal of the planned revisions to the standard will be to change the standard from a guide into a practice to provide additional information. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

12. "4.2.2 Selected units shall be analyzed to meet Practice E2329 if statistical inferences are to be made about the whole population."

    ASTM E2329 describes minimum requirements for qualitative analysis. It is not obvious why these requirements for identifying compounds do not apply to the analysis of every unit in a census of the population. Moreover, this provision seems beyond the scope of a standard on sampling. It concerns measurements or observations on the items in a sample.

Subcommittee Response: Not persuasive. Not all forensic laboratories follow E2329, and one of the references cited in E2548 discusses qualitative testing to identify a substance that would not meet E2329 requirements, but still go on to discuss making an inference to the population regarding the probability of the population containing that substance. This clause of the standard is necessary to ensure laboratories following only E2548 do test any units selected via the sampling plan according to E2329 prior to making an inference to the population regarding the probability of the population containing a substance identified in the selected samples.

13. **FIG. 1**. Relationship of Various Levels Required in Sampling



a. Previously, the leaves in the top “statistical” branch were "Hypergeometric," "Bayesian," and "Other probability-based approaches." These have been trimmed to “Bayesian” and “Frequentist.” This is an improvement, since “Hypergeometric” is merely the name of a probability distribution, and not a form of sampling. (When sampling without replacement, the probability that a number *X* of units with a discrete feature will be picked in *N* draws follows the hypergeometric distribution. This is true for Bayesians and frequentists alike.)

Subcommittee Response: Agreed; this improvement is one of the reasons we feel it is important to replace the superseded E2548-11e1 document with the E2548-16 version on the Registry.

b. The problem that remains in this branch is that neither “Bayesian” nor “frequentist” describes a sampling plan. For example, the results for a simple random sample (SRS) can be analyzed from either perspective, and the sample size can be pre-established according to computations with either methodology (or via likelihood computations, for that matter). Presumably, a sampling procedure that implements the SRS design would be a “sampling plan.” But the SRS design itself is neither Bayesian nor frequentist. In addition, there is a third school of statistical thought, likelihood theory, that could be used in analyzing sample data.

Subcommittee Response: Persuasive. We recognize that terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that

we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

c. The leaves of the "Non-Statistical" branch are "Square Root N," "Management Directive," and "Judicial Requirements." The first leaf refers to a procedure for establishing the size of a sample. The square root of *N* assuredly is a statistic, and unless the population size exceeds 9, the method is to draw an SRS (see United Nations Office on Drugs and Crime, Recommended Methods for Testing Opium, Morphine and Heroin: Manual for Use by National Drug Testing Laboratories, p.21). Consequently, the "square root method" can be used to obtain a sample statistic that is a statistically sound estimator of a population parameter.

   Subcommittee Response: Persuasive. We recognize that terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

d. The next two leaves can apply to any choice of a sampling design that comes from a manager or a court. The design could be statistically optimal or wasteful, but its statistical properties do not make it "non-statistical." Perhaps the branch is supposed to be "not suitable for statistical inference," but that does not lead to the three leaves either.

   Subcommittee Response: Persuasive. We recognize that terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

e. In sum, Figure 1 muddies the waters. It should be abandoned or redrawn after more appropriate and better defined terminology is in place.

Subcommittee Response: Persuasive. We recognize that terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

14. "5.2.1 The population determination shall take into account all typical forms and quantities in which exhibits may appear."
    a. What is an "exhibit"? The term appears nowhere else in the Guide.

    Subcommittee Response: Editorial - does not critically impact the interpretation of the standard. This term is commonly used within seized drugs forensic laboratories, however, it should be defined in this standard. A terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

    b. Why should atypical forms and quantities (of what?) not be considered along with typical ones?

    Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

15. "5.2.2 A population can consist of a single unit or multiple units. 5.2.3 A multiple unit population shall consist of items that are similar in relevant visual characteristics (for example, size, color, shape, etc.)."

    The wording suggests that material packaged in containers of different sizes or color is not a population of legal interest. In this form, the material is neither a "single unit" nor a "multiple unit" population. The statistical question may concern a feature of all the material, and that collection would be the population. It appears that in this standard, a "population" is a sampling frame that is constructed from the population. The literature on attribute and variable sampling as well as multistage sampling should be consulted. In sampling theory, a "sampling unit" is not necessarily a single element of the population. There are smaller units of elements at each stage.

    Subcommittee Response: Persuasive, although the change is not critical to the use of the standard as written. Typically units that are similar in relevant visual characteristics would be grouped together into subpopulations for sampling in seized drugs. We recognize that this step may not be clear based on the current wording in the standard. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

16. "5.4 When a single unit or bulk population is to be analyzed, the issue of homogeneity shall be addressed within the sampling plan."

    Now it appears that there is a third kind of population—a "bulk population" that is neither a single unit nor a combination of units. The definitions of all these populations should be in a definitions section.

    Subcommittee Response: Editorial. These terms are used within seized drugs forensic laboratories, however they should be defined to explain how they apply to this standard. A terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

17. "5.4.1 One sample is sufficient if the bulk material is homogeneous. Analysts can make bulk material homogeneous."

    Normally, a sample is the subset of elements drawn from the sampling frame that represents the population of interest. Of course, "sample" has another meaning—a single chunk of material for instrumental or visual analysis—but that is not necessarily a "sample" as the term is used in sampling theory. The sample from the population can consist of many items. Substituting "specimen" for "sample" in this section would avoid possible confusion.

    Subcommittee Response: Editorial - does not critically impact the interpretation of the standard. We recognize the terminology is an issue within the standard that we agree needs to be addressed. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.he term

18. "5.5 For a multiple unit population, the sampling plan may be statistical or non-statistical."

    See supra comment on Figure 1.

    Subcommittee Response: Persuasive. We recognize the terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

19. "5.5.1 Statistical approaches are applicable when inferences are made about the whole population. For example:
    The probability that a given percentage of the population contains the drug of interest or is positive for a given characteristic.
    The total net weight of the population is to be extrapolated from the weight of a sample."

a. Normally one would speak of the population percentage as a parameter whose value is being estimated from sample data. The estimate is an inference. A hypothesis about the true value would be another inferential statement. The "probability that a given percentage of the population contains the drug of interest or is positive for a given characteristic," if based on the sample data and a prior distribution, would be called a posterior probability for the inference. The posterior probability cannot be computed with frequentist methods for inference, and this fact should be made clear in the Standard Guide (and included in reports that contain coverage probabilities).

Subcommittee Response: Persuasive. However, this particular ASTM standard is a standard guide. Per ASTM requirements, guides are a compendium of information or series of options that do not recommend a specific course of action. We do agree that the standard could be improved and provide additional guidance to the reader and the goal of the planned revisions to the standard will be to change the standard from a guide into a practice to provide additional information.

b. The second example is not the inference; it is the objective of the study. An example of an inference would be "Extrapolating from the sample data, the total net weight of a controlled substance in the population is 500 grams."

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

20. "5.5.1.1 Published examples of statistical approaches involving general considerations:
(1) Practice E105.
(2) Practice E141.
(3) Terminology E1732."

A glossary or dictionary of terms is not an example of an approach.

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

21. (4) Guidelines on Representative Drug Sampling.[4]
[4] 2009 UNODC/ENFSI Guidelines on Representative Drug Sampling, European Network of Forensic Science Institutes (ENFSI), 2004, http://www.ENFSI.org.

A 2009 document cannot be published in 2004.

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

22. (5) ISO 3534-1.
(6) ISO 3534-2.

A glossary or dictionary of terms is not an example of an approach, but maybe some of the definitions contain such examples.

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

23. "5.5.2 Non-statistical approaches are appropriate if no inference is to be made about the whole population.

5.5.2.1 Published examples: (1) The "square root method."[7] (2) Methods listed in "Arbitrary Sampling" in *Guidelines on Representative Drug Sampling*."

The "square root method" of (1) is a form of "arbitrary sampling" according to the *Guidelines* cited in (2). Therefore, having the two categories of "non-statistical approaches" makes little sense (even if one can construe using an SRS with the various "square root methods" listed in the EU document for choosing the sample size as "non-statistical").

Subcommittee Response: Persuasive. We recognize the terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

24. "5.5.2.3 A non-statistical sampling approach may allow an inference about the population. If a single population has been randomly sampled, the data may allow an inference to be drawn by (1) determining and reporting a confidence interval for an inferred population parameter (for example: weight or tablet count); or by (2) retrospectively using the results in a statistical model and determining the resulting probabilities and level of confidence."

This paragraph is a real puzzler. The essence of probability sampling is that the probability that each and every element in the sampling frame is known. Knowing these probabilities allows one to use a statistical model to determine the sampling distribution of the estimator. That, in turn, leads to a standard error, a point estimate, and thus a confidence interval. If "non-statistical sampling" does not use probability sampling, how is one supposed to determine the standard error? The answer might be to analogize the approach actually used to a probability sampling method of some sort. But that interpretation of the paragraph is contradicted by "(2) retrospectively using the results in a statistical model and determining the resulting probabilities and level of confidence." (1) is already retrospective and uses a statistical model to produce the confidence interval, so what is (2) supposed to add? Furthermore, (2) is obscure at best. A "level of confidence" is a coverage probability. As such, what are "the resulting probabilities"?

Subcommittee Response: Persuasive. We recognize the terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

25. "5.6.1 Establish the procedure for selecting the number of units that will comprise a sample."

    This sounds prescriptive, but the relative merits of different procedures have, strangely, not been discussed in this Standard Guide. Accordingly, the document supplies no guidance beyond a general admonition to read the literature mentioned in it and to do the right thing. What would that be? The UN report, for example, contends that even though the square-root sample-size determination lacks a sound theoretical foundation, it gives reliable results. The EU report, which describes this rule-of-thumb as "arbitrary," acknowledges that it does "work well in many situations," and explicates a number of variations on it. But statisticians have been known to dismiss it out of hand. E.g., Alan Julian Izenman, Statistical and Legal Aspects of the Forensic Study of Illicit Drugs, 16 Stat. Sci. 35–57, 47 (2001) ("The popularity of the square root rule, despite the lack of theoretical support for this rule, shows how an unfounded rule-of-thumb can be established in the practice of a particular field."). What guidance does the Standard Guide have to offer?

    Subcommittee Response: Not persuasive as to this version, but will be considered as part of the full revision of the standard. This particular ASTM standard is a standard guide. Per ASTM requirements, guides are a compendium of information or series of options that do not recommend a specific course of action. We do agree that the standard could be improved and provide additional guidance to the reader and the goal of the planned revisions to the standard will be to change the standard from a guide into a practice to provide additional information.

26. "5.6.1.1 For non-statistical approaches, select a sample appropriate for the analytical objectives."

    There is no indication in the Standard Guide of how to go about doing this and no clear statement of what "analytical objectives" are. Usually, "analysis" in the Standard Guide refers to laboratory analysis (particularly § 6), but in this subsection the phrase "analytical objectives" seems to be a veiled reference to doing what management or the law dictates. Or, if that is not what the words point to, does not the same advice apply to "statistical approaches"? Surely, in all situations, one should select "a sample appropriate for the analytical objectives." The reader is left uninformed by the Guide.

    Subcommittee Response: Not persuasive as to this version, but will be considered as part of the full revision of the standard. This particular ASTM standard is a standard guide. Per ASTM requirements, guides are a compendium of information or series of options that do not recommend a specific course of action. We do agree that the standard could be improved and provide additional guidance to the reader and the goal of the planned revisions to the standard will be to change the standard from a guide into a practice to provide additional information.

27. "5.6.1.2 For statistical approaches, random sampling shall be conducted."

    Given that probability sampling always is possible in this context (at least approximately), why would one ever want anything less? For example, if the sample size n is determined by the formula $n = N/10$, where N is the population size (based on an opinion in which some judge at some time said that he or she thought that samples of less than 10% could not be trusted), are we in the 5.6.1.1 realm of "non-statistical approaches"? The selection method should use a randomizing mechanism regardless of a "non-statistical" approach to determining the sample size.

Subcommittee Response: Not persuasive as to this version, but will be considered as part of the full revision of the standard. Sample selection up to a weight threshold may not require random selection since no inference to the population is made. We recognize the terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

28. "5.6.2.1 A random sample is one selected without bias and where each item has an equal chance of being selected."

    The words "without bias" seem like they represent a separate requirement, but they are merely superfluous. The reason to give an item in a sampling unit (which could be the sampling frame) an equal probability of being selected is to avoid bias (and give estimates of sampling error a solid foundation). The sentence could be broken into the three different thoughts: "A simple random sample is one in which each item in the sampling frame has an equal probability of being selected. Using a randomizing procedure to equalize the probability avoids selection bias and supplies the foundation for estimating sampling error."

    Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

29. Why is simple random sampling the only acceptable form of probability sampling?

    Subcommittee Response: Not persuasive as to this version, but will be considered as part of the full revision of the standard. There is no negative impact to the standard by only discussing random sampling. The addition of other forms of probability sampling may be considered during the next revision to the standard.

30. "5.6.2.1 Computer generated random numbers or random number tables are commonly employed for such tasks and these should be included in the sampling plan.

    5.6.2.2 Random sampling of items using random number tables may not be practical in all cases. In these instances, an alternate sampling plan shall be designed and documented to approach random selection. A practical solution involves a "black box" method, which refers to one that will prevent the sampler from consciously selecting a specific item from the population (that is, all units are placed in a box and the samples for testing are selected without bias)."

    5.6.2.1 requires a sampling plan that provides for selection according to random numbers even when their use is impractical. Then 5.6.2.2 says not to use that part of the plan in some cases. The final clause of 5.6.2.1 should be "and these should be part of the sampling plan unless a statement showing that they are impractical is included in the written plan."

Subcommittee Response: Not persuasive. 5.6.2.1 states that random number generators should be used, they are not required. 5.6.2.2 provides guidance on what to do when a random number generator is not able to be used.

31. Why are "random number tables" but not pseudo-random number generators the subject of 5.6.2.2?

    Subcommittee Response: Not persuasive. There is no negative impact to the standard by only discussing random number tables. The addition of other forms of random number generators may be considered during the next revision to the standard.

32. "5.7 Sample Reduction—Sample reduction may be applied in cases where the weight or volume of the selected units is too large for laboratory analysis (Fig. 2, insert A)."

    "Sample reduction" should be defined. Figure 2 does it to some extent (by listing the names of three methods for reducing the quantity), but a section with a general definition would improve readability. Again, in the figure, "sample" is used in the laboratory-analysis sense rather than the sampling-from-a-population sense. A word or phrase like "specimen" or "analytical unit" would be better for the former kind of sample.

    Subcommittee Response: Editorial - does not critically impact the interpretation of the standard. This term is commonly used within seized drugs forensic laboratories, however, it should be defined in this standard. A terminology section of terms specific to the standard would be beneficial and strengthen the standard; the planned revisions to the standard will incorporate a terminology section.

33. "6. Analysis
    6.1 Statistically Selected Sample(s)—In accordance with 4.2.2, it is recommended that each unit comprising the sample be analyzed to meet Practice E2329 if statistical inferences are to be made about the whole population.
    6.2 Non-statistically Selected Sample(s)—Practice E2329 shall be applied to at least one unit of the sample."

    See comment on § 4.2.2.

    Subcommittee Response: Previously considered in reference to 4.2.2.

34. "7. Documentation
    7.1 Inferences drawn from the application of the sampling plan and subsequent analyses shall be documented."

    This section is exceedingly skimpy. It does not even cover the documentation mentioned in § 4.2.1.2 and § 5.6.2.2

    Subcommittee Response: Persuasive. Sections 4.2.1.2 and 5.6.2.2 do provide some additional requirements on documentation and reporting and there are separate ASTM reporting standards. However, we agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the

OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

35. "8. Reporting
8.1 Sampling information shall be included in reports."

What "sampling information" should be reported? The subsections that follow (8.1.1 and 8.1.2) refer to nothing more than a statement of the results. They do not discuss reporting the parts of the sample design and selection mechanisms that led to the testing.

Subcommittee Response: Persuasive. Sections 4.2.1.2 and 5.6.2.2 do provide some additional requirements on documentation and reporting and there are separate ASTM reporting standards. However, we agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

36. "8.1.1 Statistically Selected Sample(s)—Reporting statistical inferences for a population is acceptable when testing is performed on the statistically selected units as stated in 6.1 above. The language in the report must make it clear to the reader that the results are based on a probability-based sampling plan. 8.1.2 Non-Statistically Selected Sample(s)—The language in the report must make it clear to the reader that the results apply to only the tested units. For example, 2 of 100 bags were analyzed and found to contain Cocaine."
    a. The Guide is confusing in its description of what qualifies as "a probability-based sampling plan." For example, it appears from Figure 1 and parts of the text noted above that if the sample size is not chosen by using probability theory at that stage, the result is a "non-statistically selected sample." The report therefore could not make any extrapolation to the population from an SRS with an ad hoc sample size. But using an unnecessarily large sample—which is what using a statistically optimal sample size guards against—does not undermine the validity of a statistical inference from the SRS. The laboratory should be able to make inferences from probability samples regardless of the method for choosing the sample size.

    Subcommittee Response: Persuasive. We recognize the terminology and the distinction between sample selection and sampling is a fundamental issue within the standard that we agree needs to be addressed. We understand that the terms "statistical and non-statistical sampling" may be unclear to some readers and may be interpreted differently by various disciplines. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

b. When the sample is not a probability sample—one drawn so as to give every element in the sampling frame a known probability of selection—an ampliative inference is of questionable validity, and the demand that the report "make it clear to the reader that the results apply to only the tested units" is reasonable. But merely stating that *x* out of *N* bags "were analyzed and found to contain Cocaine" does not make it clear that the sample proportion *x*/*N* might deviate greatly from the population proportion and that, without simple random sampling, statistical theory cannot be used to estimate the sampling error. Moreover, although the "*x* of *N*" format for reporting is just as appropriate for estimates from probability samples, § 8.1.1 does not mention it. It only addresses the need to report that "a probability-based sampling plan" was used. It overlooks the issue of how to report the results themselves.

Subcommittee Response: Persuasive. While there are separate ASTM reporting standards, we agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

37. "9. Keywords
9.1 analytical method; qualitative method; quantitative method; sampling; seized drug analytical method"

The only keyword that has to do with sampling is "sampling." For a Standard Guide on sampling, that is shocking.

Subcommittee Response: Editorial - does not critically impact the interpretation of the standard.

Commenter 2:

E2548 Standard Guide for Sampling Seized Drugs for Qualitative and Quantitative Legal Task Group Comments. The primary role of the Legal Task Group is to review standards with an eye towards how the standards will be used in the legal system. Sections of standards addressing documentation, reporting, and testimony are therefore of great concern to this task group. For the reasons set forth in these comments, we do not believe this standard belongs in the Registry.

1. Inadequate reporting which allows for testimony that overstates the conclusions reached by the analyst has been one of the leading causes of wrongful convictions in the area of forensic science. It is essential that consensus standards advise laboratories to establish protocols that require analysts to clearly and fully document files so that another analyst examining the file may see precisely what the analyst did and how she came to her conclusion. Reports should be similarly detailed so that the end user of the report is apprised of the conclusions, the basis for the conclusions and their limitations. The National Commission on Forensic Science expressed the view "that Forensic Science Service Providers (FSSPs) and Forensic Medicine Service Providers

(FMSPs) should have written policies for documenting the examination, testing, or interpretation of evidence and for the reporting of results, interpretations, and conclusions that are consistent with the following requirements:

a. Records should be created during the examination of evidence and during the technical review that would allow another analyst or scientist with proper training and experience to understand and evaluate all the work performed and to independently analyze and interpret the data and draw conclusions.
b. Records created by FSSPs and FMSPs should also provide information necessary for use in the criminal justice system (e.g., chain of custody).
c. Providing all of the documentation described above in a single report in every case is impractical. Instead, if not in the report, the documentation described herein must be maintained in a case record if it is specific to a c///se or test. Generic documentation such as standard operating procedures and definitions must either be a part of the case record or be easily accessible (e.g., posted on a web site, made available on request).
d. Reports should accurately and clearly convey a statement of the purpose of the examination, testing, and interpretation of the evidence; the method and materials used; a summary or a description of the data or results obtained; any conclusions or interpretations derived from the data or results; any discordant results, interpretations, or conclusions; and, where necessary for the interpretation of test results, sources of uncertainty in the procedure and conclusions along with estimates of their scale.
e. Every report should include a statement that the report does not contain all of the documentation associated with the work performed or what is necessary to understand and evaluate all the work performed, and to independently analyze and interpret the data and draw conclusions requires a review of the case record. The case record should be organized; and made available in a manner consistent with the discovery Recommendations of the National Commission on Forensic Science."

E2548 fails to give adequate guidance regarding what must be documented and reported and thus fails to provide labs the guidance the criminal justice system requires.

Subcommittee Response: Persuasive. While there are separate ASTM reporting standards, we agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

2. "7. Documentation 7.1 Inferences drawn from the application of a sampling plan and subsequent analysis shall be documented."

This section (or some other) must specify the essential elements of the sampling plan that are subject to this documentation requirement. Plainly, it is not sufficient to document only the

inferences. The meaning of "subsequent analysis" is obscure. Is it the tests used to identify or quantify the seized material? A confidence interval or some other statistic that indicates sampling uncertainty? Something else?

Subcommittee Response: Persuasive. We agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

3. "8. Reporting 8.1 Sampling information shall be included in a report."
Neither section provides the requisite level of specificity. As written, the standard does not require labs to provide legally sufficient documentation or reporting. At a minimum the sampling plan and how it was chosen should be documented. The limits of the plan that was used should be clearly reported. An analyst reviewing the file should know precisely what plan was used, why it was selected and the limits of the conclusion reached given the plan chosen.

Subcommittee Response: Persuasive. While there are separate ASTM reporting standards, we agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

4. Section 8.1.1 states that "The language in the report must make it clear to the reader that the results are based on a probability-based sampling Plan." To say that sampling involved some unspecified randomizing mechanism (which is what we presume "probability-based" means) is grossly incomplete. The report should state which sampling plan was actually used. If the sampling as performed deviated from the plan adopted in advance, that should be stated and justified. The report should state the sampling error or include some other measure of uncertainty arising from sampling — as well as from other sources. In short, the standard is missing any reasonably complete statement or explanation of what to include in a report about a sampling plan.

Subcommittee Response: Persuasive. While there are separate ASTM reporting and uncertainty standards, we agree that this section should be expanded and an uncertainty section added. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

5. Section 8.1.2 states that for "Non-Statistically Selected Sample(s) —The language in the report must make it clear to the reader that the results apply to only the tested units. For example, 2 of 100 bags were analyzed and found to contain Cocaine." This example does not make it clear "that the results apply to only the tested units." It is simply a statement of the results. Whenever sampling is used, such a statement of results should be in the report. The standard is missing any reasonably complete explanation of what to include in a report about the limits of what it calls "non-statistical sampling."

   Subcommittee Response: Persuasive. While there are separate ASTM reporting standards, we agree that this section should be expanded. To that end, a work item has been opened at ASTM (WK75229) and the OSAC Seized Drugs Subcommittee will form a task group to revise the document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 3+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry.

OSAC Registry Request Comment Adjudication Response Template

Document Title	ASTM E2548-16: Standard Guide for Sampling Seized Drugs for Qualitative and Quantitative Analysis (ASTM International)
Requesting Unit	Seized Drugs SC

Unit Chair

Name: Agnes Winokur
Affiliation: DEA

Unit Technical Contact

Name: Anne Slaymaker
Affiliation DEA

Beginning Comment Period Date	10/6/20	
End Comment Period Date	11/6/20	
Comment Adjudication Meeting Dates	12/8/20	12/14/20
# of Members Present	5	4
Resolution Date and Vote Outcome	16 YES, 0 NO, 1 Abstain; Ballot Closed: 02/26/2021 03/01/2021 - public comment adjudication approved by OSAC Sezied Drugs Subcommittee	

Note: This template is intended for use by all Units considering a new document for addition into OSAC Registry

Area
ience

ID	Person	Section	Page	Line	Editorial or Technical	Attached File
1	David H Kaye					ASTME2548-16-K-201106
2	Jennifer Friedman					none
3	Ramon O Diaz					none

Comment/Proposals
Neither ASTM E2548-11e1 nor its successor, ASTM E2548-16, belong on "a repository of high-quality, technically sound pub standards for forensic science." The basis for this conclusion can be found in the attached memorandum. If the choice is bet ASTM E2548-11e1 to remain on the Registry and replacing it with ASTM E2548-16 pending a major overhaul, however, the p question of whether the 2016 version belongs on the Registry is more complex.
E2548 Standard Guide for Sampling Seized Drugs for Qualitative and Quantitative Legal Task Group Comments The primar Legal Task Group is to review standards with an eye towards how the standards will be used in the legal system. Sections of st addressing documentation, reporting, and testimony are therefore of great concern to this task group. For the reasons set fo comments, we do not believe this standard belongs in the Registry. Inadequate reporting which allows for testimony that conclusions reached by the analyst has been one of the leading causes of wrongful convictions in the area of forensic science that consensus standards advise laboratories to establish protocols that require analysts to clearly and fully document files s analyst examining the file may see precisely what the analyst did and how she came to her conclusion. Reports should be sim that the end user of the report is apprised of the conclusions, the basis for the conclusions and their limitations. The Natio on Forensic Science expressed the view "that Forensic Science Service Providers (FSSPs) and Forensic Medicine Service Provi should have written policies for documenting the examination, testing, or interpretation of evidence and for the reporting interpretations, and conclusions that are consistent with the following requirements: 1. Records should be created during of evidence and during the technical review that would allow another analyst or scientist with proper training and experien and evaluate all the work performed and to independently analyze and interpret the data and draw conclusions. 2. Records and FMSPs should also provide information necessary for use in the criminal justice system (e.g., chain of custody). 3. Provi documentation described above in a single report in every case is impractical. Instead, if not in the report, the documentati herein must be maintained in a case record if it is specific to a c///se or test. Generic documentation such as standard operat and definitions must either be a part of the case record or be easily accessible (e.g., posted on a web site, made available on r Reports should accurately and clearly convey a statement of the purpose of the examination, testing, and interpretation of t method and materials used; a summary or a description of the data or results obtained; any conclusions or interpretations c data or results; any discordant results, interpretations, or conclusions; and, where necessary for the interpretation of test re uncertainty in the procedure and conclusions along with estimates of their scale. 5. Every report should include a statemen does not contain all of the documentation associated with the work performed or what is necessary to understand and eval performed, and to independently analyze and interpret the data and draw conclusions requires a review of the case record. should be organized; and made available in a manner consistent with the discovery Recommendations of the National Comm Forensic Science." E2548 fails to give adequate guidance regarding what must be documented and reported and thus fails t the guidance the criminal justice system requires. 7 Documentation Inferences drawn from the application of a sampling pl subsequent analysis shall be documented. This section (or some other) must specify the essential elements of the sampling p subject to this documentation requirement. Plainly, it is not sufficient to document only the inferences. The meaning of "su analysis" is obscure. Is it the tests used to identify or quantify the seized material? A confidence interval or some other statis It should be placed on the OSAC Registry because it will improve the availability to forensic drug chemist and laboratories.

Response
ASTM E2548-11e1 “Standard Guide For Sampling Seized Drugs for Qualitative and Quantitative Analysis” is currently on the OSAC Registry. This standard has been superseded at ASTM by E2548-16; as such, the OSAC Seized Drugs Subcommittee submits ASTM E2548-16 to replace the historical version of the standard on the OSAC Registry. We feel it is important to replace the superseded E2548-11e1 document with the E2548-16 version on the Registry since several improvements were made in the E2548-16 document and it is in the best interest of the forensic community for OSAC to refer individuals to the current version of the document. The improvements made in E2548-16 include: (1) requiring random sampling when making an inference to a population, (2) providing information on the possiblity of making an inference to the population from non-statistical sampling plans, and (3) improvements to Figure 1. We do agree with several of the public comments and support a full revision of the standard. To that end, a work item has been opened at ASTM and the seized drugs subcommittee will form a task group and invite a statistician and an LRC member (especially Chairperson David Kaye, who has provided many persuasive comments) into the task group to participate in the drafting of the revised document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 2+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry. **See specific comment adjudication in the attached document.**

ASTM E2548-11e1 “Standard Guide For Sampling Seized Drugs for Qualitative and Quantitative Analysis” is currently on the OSAC Registry. This standard has been superseded at ASTM by E2548-16; as such, the OSAC Seized Drugs Subcommittee submits ASTM E2548-16 to replace the historical version of the standard on the OSAC Registry. We feel it is important to replace the superseded E2548-11e1 document with the E2548-16 version on the Registry since several improvements were made in the E2548-16 document and it is in the best interest of the forensic community for OSAC to refer individuals to the current version of the document. The improvements made in E2548-16 include: (1) requiring random sampling when making an inference to a population, (2) providing information on the possiblity of making an inference to the population from non-statistical sampling plans, and (3) improvements to Figure 1. We do agree with several of the public comments and support a full revision of the standard. To that end, a work item has been opened at ASTM and the seized drugs subcommittee will form a task group and invite a statistician and an LRC member (especially Chairperson David Kaye, who has provided many persuasive comments) into the task group to participate in the drafting of the revised document. The challenge faced is that the process of revising the standard through the OSAC and ASTM processes will take 2+ years and we feel the addition of the compulsory language regarding random sampling is an important enough change to support replacement of E2548-11e1 with E2548-16 on the Registry. **See specific comment adjudication in the attached document.**

No response needed.

Reason	Resolution
	Persuasive - review required
	Persuasive - review required
	No response needed

No response needed

Categories for adjudication of negative public comment for addition to registry

Term	Definition
Not Germane	Comment is not relevant to the subject of document being considered
Persuasive - review required	General agreement with comment given, further review by subcommittee
Withdrawn by submitter	Comment withdrawn by submitter
Not persuasive	Justification for non persuasive rationale is indicated by committee action
Previously considered	Topic of comment was previously discussed and resolved by subcommittee*
No response needed	Comment does not require a response

*If all commenters express the same or similar reasoning for not moving a document forward, the OPO looks to confirm that all similar comments have the same adjudication response (e.g. Persuasive, Not Persuasive, etc.). We also look to confirm that the response provided by the unit is the same or similar. This is to ensure that when the final comment(s) and response(s) are published to the OSAC website, that viewers have a clear understanding of the unit's perspective on a given issue.