AdvaMedDx
Vital Insights | Transforming Care

The Coalition for CLIA Waiver Reform

AdvaMedDx and Coalition for CLIA Waiver Reform Proposal: Proposed Section of Guidance for Assessing Whether a Test is "Accurate" for CLIA Waiver Purposes (November 2015)

1. Purpose of Guidance

The purpose of this guidance is to present FDA's current thinking on how to establish that an *in vitro* diagnostic ("IVD") is accurate for the purposes of CLIA waiver applications.

1.1 Status of 2008 CLIA Waiver Guidance

This document (or otherwise referred to as "guidance") provides proposed revised guidance to replace Section V of "Guidance for Industry and FDA Staff: Recommendations for Clinical Laboratory Improvement Amendments of 1988 (CLIA) Waiver Applications for Manufacturers of In Vitro Diagnostic Devices," issued on January 30, 2008. This guidance sets forth updated provisions to support demonstration of whether a test is "accurate" for CLIA waiver purposes, and provides different options for evaluating whether a test is "accurate" for CLIA waiver purposes under the Agency's current interpretation of accuracy. Sponsors are encouraged to discuss alternative approaches, or hybrids of these approaches, with FDA that are appropriate for their particular technology and the IVD's intended use.

Other Sections of the January 30, 2008 Guidance continue to represent FDA's current thinking.

1.2 FDA's Interpretation of "Accurate" Under 42 U.S.C. § 263a(d)(3)(A)

There are various pathways by which a diagnostic test can be granted a CLIA waiver. One pathway, established by 42 U.S.C. § 263a(d)(3)(A), is a determination by FDA that the tests "employ methodologies that are so simple and accurate as to render the likelihood of erroneous results by the user negligible." The "by the user" language was added to clarify the assessment for accuracy (for CLIA waiver purposes) –

> "should focus on the test performance "by the user" and the potential for operator error in performing the test. The purpose of CLIA quality control, proficiency testing, and personnel requirements is to ensure consistent, reliable, and appropriate use of a test system by users of the test. Without the clarifying

> "by the user," interpretations of "erroneous results" and "accurate" could include the inherent clinical sensitivity/specificity of a test system, parameters that are properly reviewed [in] determining whether to approve of or clear a product for marketing."[1]

Consistent with this law, FDA interprets "accurate" in the context of CLIA waivers to mean that the skill of the user does not have a meaningful impact on results obtained under intended operating conditions (i.e. as used in Certificate of Waiver testing facilities at the point-of-care). Thus, the purpose is to evaluate the effect that the "user" has on test results, not to revalidate previous determinations with respect to safety and effectiveness. Any study design and analysis which shows that the skill of the user does not meaningfully impact test results is sufficient to establish accuracy for a CLIA waiver application. Therefore, other study designs that are not discussed in this guidance might be considered, especially if unique needs or limitations occur.

2. Study Design and Statistical Considerations

2.1 Number of Users

The objective of the studies described herein is to assess whether specialized laboratory training and experience of operators found in moderate and high complexity test settings, including those medical personnel who perform tests in moderate or high complexity point of care settings (collectively, "Trained Users") achieve the same results as operators that are expected to be found at Certificate of Waiver sites (referred to herein as "Untrained Users").[2] Therefore, the number of operators should be sufficient to allow for a meaningful analysis that captures the variability of skills within both user populations.[3]

2.2 Demographics of Participants

Untrained Users in the study should include anticipated test users. These users may include personnel who have no specific formal training in performing in vitro diagnostic tests, for example, physicians, nurses, medical or nursing assistants, and/or others who would be found in the intended testing environment and be responsible for testing. It is not necessary that these users have no experience with tests generally, only that they are representative of the

[1] H.R. Rep. No. 105-310, Sec. 21.

[2] Although referred to as "untrained users," because they have no specific training in operation of the IVD for which a waiver is being sought, we note that the vast majority of these users are credentialed medical professionals who understand the importance of diagnostics test quality and may have experience with other tests.

[3] Clinical and Laboratory Standards Institutes (CLSI), *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline--Second Edition.* CLSI document EP12-A2 (ISBN 1-56238-654-9). Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, PA 19087-1898 USA, 2008.

intended user population (most of whom would likely have some experience with testing). Demographic information, including occupation, highest level of education and experience level should be collected as part of the trial.

2.3 Instructions for Use

Untrained Users should only be provided with whatever materials they would have access to when purchasing the test. This could include traditional instructions for use ("IFU"), a "start-up" card or "Quick Reference Guide" that explains how to use the test, or visual or audio aids (e.g., instructional videos provided with test kits) if these are included in the materials provided to customers. Materials may also include an 800-number helpline, or an internet address for on-line support, if such services are to be provided for the IVD when it is marketed. Once materials are provided to the users, they should self-train on materials, and data should be collected with a survey after their participation in the study to confirm use of materials by participants and to collect information on what materials users found most useful to using the test.

In addition, if you intend to make a "startup" kit available that contains controls or other materials for operator self-training purposes, it may be used in your study design. Those materials would be subject to other regulatory controls.

2.4 Samples and Subjects

A common approach in CLIA waiver studies is to collect fresh samples from subjects under real world conditions and presentations. For example, if a test is intended for screening unknown infections or other health conditions in persons with certain clinical presentations, the subjects in the study would reflect those patients.

If the samples from the subjects in the study do not exhibit analyte concentrations[4] that span the full measurable range of the IVD (for quantitative tests), or do not provide a sufficient number of positive and negative samples or samples near the cut-off (for qualitative tests) , then contrived or banked samples may generally be used to supplement other data. Also, in some instances, the entire set of samples used in the study may be comprised of banked or contrived samples, as discussed in Section 2.5.

2.5 Use of Fresh, Banked, and Contrived Samples

In order to receive a waiver under the standard of 21 USC § 263a(d)(3)(A), a diagnostic test must be simple to use. Among other things, this means that samples or reagents would be subject to basic, non-technique dependent manipulation. Consequently, user influence on sample-related interactions during sample collection is generally expected to be minimal.

We recognize that there are instances where the use of contrived or banked samples is a reasonable option to supplement, or entirely replace, the use of freshly-taken subject samples

[4] We use "analyte concentrations" to refer to amounts of a substance of interest in a sample, be it the amount of a chemical, biological material, living organism (such as colony forming units), or any other measurand.

in a waiver evaluation. Examples of these instances include (1) situations where it might not be possible to find natural samples that span the analytical range in the settings of intended use for a waived diagnostic, (2) collection conditions are rare, such as samples from patients from rare diseases, or patients who have weak positive or negative analyte levels, or (3) the collection procedure is standard and routine in the waived setting (e.g., a venous puncture, standard finger stick or swab) and thus would not be expected to be a determinant in analytical performance. We recommend that you include a discussion of the rationale for including contrived or banked samples in your study when such samples are proposed.

When fresh patient samples are collected for a study, the samples should be collected from subjects who reflect the population of patients who will be tested with the proposed IVD. We recommend that you include a discussion of how the data collected generally correspond to populations that are most likely to be tested with the IVD.

Finally, for products that can be used with multiple sample matrices,, CLIA waiver accuracy studies on multiple sample matricies may be unnecessary to support a waiver in all of those matrices. The effects of sample matrices on accuracies are assessed as part of a 510(k) or PMA premarket review. A CLIA waiver accuracy study, on the other hand, is intended to assess the effects that variation in user expertise may have on the accuracy of results. So, for example, when the steps in analysis of two different matrices with the IVD are essentially the same, the impact of user expertise on the accuracy of analyzing each matrix will be identical, and there would generally be no need to conduct studies with both matrices. Please consult FDA if there are questions regarding necessary sample types for CLIA waiver studies. .

2.6 Additional Study Design Considerations

The fundamental question that a CLIA waiver accuracy study is intended to answer is if the results obtained by Trained Users and Untrained Users are equivalent. If so, then the test is "accurate" for CLIA waiver purposes. In addition, there may be situations where a difference in results obtained by Trained Users and Untrained Users may exist, but the difference is (1) not considered clinically meaningful, and/or (2) is considered acceptable because the benefits of expanding access to point-of-care testing outweigh risks that result from a difference in performance.[5]

With this in mind, this guidance provides approaches for assessing comparability based on measures of accuracy and agreement, respectively.

[5] It is important to note that as technologies improve and IVDs become more precise, it becomes more easily apparent to identify differences, even when those differenced may not be meaningful. Therefore, where differences do occur, they need to be assessed against the needs of the physician and patient.

2.6.1. Accuracy

A test may be shown to be accurate for CLIA waiver purposes if a set of appropriate and representative samples with assigned values are analyzed by both a group of Trained Users and a group of Untrained Users, and the two groups achieve equivalent accuracy. The individual elements of accuracy – sensitivity and specificity – should be evaluated as they generally have different clinical relevance, and equivalence criteria should be developed for each taking the risk-benefit and test performance factors described in 2.6.3 into account.

There also might be instances where Trained User performance data in prior 510(k) or PMA submissions can be used as a historical control. If sufficiently similar sample sets are used,[6] the accuracy of a group of Untrained Users is compared to the accuracy obtained by Trained Users that was observed in studies from a 510(k) or PMA submission. The data sets in 510(k) and PMA submissions addressing inter-operator performance of Trained Users may be particularly useful if this option is appropriate and used. In this case, only Untrained Users would participate in the CLIA waiver study, but sufficient information must be provided to establish that variation in the time, samples used and conditions under which the study was performed (beyond the difference in users being studied) will not affect the interpretation of results.

2.6.2. Agreement

A test may also, in some cases, be shown to be "accurate" for CLIA waiver purposes if Trained Users and Untrained Users achieve a sufficient level of agreement. Although agreement studies do not provide a direct comparison of test accuracies in the hands of Trained Users and Untrained Users, they can be used to assess the frequency with which Trained Users and Untrained Users will get equivalent results, potentially providing reasonable assurance that the level of user expertise does not affect test results.

For this approach to be used, it is necessary to determine the degree to which results must agree (and, for quantitative tests, the amount of variation between Trained User and Untrained User groups that would be considered to agree). In some instances, during the 510(k) or PMA process, FDA has, in effect, agreed to criteria developed by the sponsor that reflects acceptable inter-operator performance amongst Trained Users. In those cases, a sponsor may rely upon the same criteria for Untrained Users for the purposes of a CLIA waiver application.

Agreement intervals may be developed independently taking into consideration both the reproducibility of the test when performed by Trained Users and the use of the test for clinical purposes, and other relevant factors as described in Section 2.6.3.

[6] Sufficient similarity is a matter of scientific judgment, and we recommend discussing this approach with FDA prior to initiating studies. In some instances sufficient similarity to allow for a cross-study comparison of results would require the use of banked samples from the original study. In other instances, separate samples sets would be acceptable provided the demographics of subjects and essential properties of samples collected were comparable.

2.6.3. Levels of Accuracy or Agreement That Should Be Achieved to Satisfy CLIA Waiver Requirements

What constitutes equivalence of Trained User and Untrained User sensitivities and specificities, or the minimum level of agreement that must be achieved between Trained User and Untrained User groups, can vary from test to test. In general, equivalence criteria and minimum levels of agreement should be determined based on risk/benefit and test performance factors, including but not limited to: (1) the clinical use of the test; (2) the clinical importance of the parameter being evaluated (e.g., sensitivity, specificity, percent positive agreement, percent negative agreement; criteria for each parameter might be different depending upon its clinical importance); (3) the role of the test in diagnosis (e.g., is the test intended to be determinative, or as an aid in diagnosis in which other clinical presentation and information is available); (4) whether confirmation is required; and (5) the sensitivity and specificity of the test when performed by Trained Users.

3. Quantitative IVD Studies

Two options for CLIA waiver accuracy evaluations of quantitative tests are discussed in this section –

1. Agreement studies using subjects enrolled in a CLIA waiver clinical trial; and
2. Equivalence of accuracy studies using samples with assigned values.

The sponsor should determine which option is most appropriate for their IVD. Also, sponsors are encouraged to discuss alternative approaches, or hybrids of these approaches, with FDA that are appropriate for their particular technology and the IVD's intended use. For example, each of the options described below assume that both Trained Users and Untrained Users will participate in the study, though as discussed in Section 2.6, there may be approaches where data or criteria from 510(k)s or PMAs may be used in lieu of Trained Users in the CLIA waiver study. In those instances, sponsors may want to discuss adopting the study designs described below to leverage the pre-existing data or criteria.

3.1 Agreement Study with Subjects

The following outlines the study design elements and statistical considerations for and agreement study using study subjects in a clinical trial.

- **Sites**. A minimum of three study sites should be included.
- **Samples**. In general, samples should be obtained fresh from subjects *representing the real-world presentation of patients*. For example, if a test will be used as an aid in diagnosis or to monitor patients with known conditions, the subjects in the study should have those attributes, as those would reflect the real-world conditions of use.

  Samples should also have sufficient variation in concentration to span the measuring range of the test unless there is a reasonable basis to conclude that only a portion of the

range should be included in the study. If samples from subjects are used in the study, but they do not exhibit concentrations spanning the applicable range, other sources of samples may be used to complete the study sample set. For example, contrived or banked samples may be used to supplement fresh samples. Another option may be to accept subjects without attributes of the indicated population, but which have the analyte concentrations needed for the study. See discussion in Section 2.5 for more information.

- **Number of Samples**. The number of samples required will vary depending on test, user variation, agreement intervals and other factors, though historically 360 samples total (approximately 120 samples per site) has been sufficient for CLIA waiver applications. If it is desirable to use more specific powering calculations, it may be helpful to use inter-operator studies with Trained Users to establish performance; these studies may be available in premarket submissions, or could be independently developed. The studies should be powered such that the sample size applies to the data collected from all sites combined; each site does not need to be powered independently.
- **Untrained Users**. The number of Untrained Users should be sufficient to ensure a representative user population, and sufficient statistical power to allow for a fulsome evaluation of the tests to a statistical significance. Generally, a study with nine Untrained Users should be sufficient, though a sponsor may choose to use more.
- **Trained Users**. If the sponsor chooses not to use data from a prior 510(k) or PMA study to establish Trained User performance, a minimum of six Trained Users should participate. However, a lesser number of users might be appropriate if results are not dependent on visual interpretation.
- **Analysis of Samples**. For each subject, samples should generally be independently collected and analyzed by any one Trained User and any one Untrained User at a study site. However, in those instances where sample collection technique is not likely to affect results (e.g., venipuncture) it may be permissible for one sample to be collected by the Trained User or Untrained User, and analyzed by both. In addition, batch testing of Trained User samples from multiple sites at a single central laboratory site may be considered.
- **Statistical Analysis**: Agreement between Trained Users and Untrained Users can be assessed by adapting the Bland-Altman method for assessing agreement between two methods in which results obtained by a Trained User and an Untrained User from the same samples are used.[7] Agreement intervals that define the bounds of sufficient agreement would be developed using the principles described in Section 2.6.2 and 2.6.3. Data analysis should be presented and considered in the aggregate across all users and

[7] See Section 5 in "*Measuring agreement in method comparison studies by Bland and Altman*," Stat Methods Med Res 1999; 8; 135-160).

sites. For additional general guidance of statistical methods for quantitative tests, please refer to CLSI EP9.

3.2 Accuracy studies using samples with assigned values

For tests that employ simple sample collection (e.g., a venous blood sample, urine sample, nasal swab, or throat swab), it is generally unnecessary to include the sample collection process as part of the study design because it can be assumed that Untrained Users will have equivalent or better specimen collection skills than Trained Users. In these cases, it is permissible to use a set of samples collected independent of the analysis step – these may be fresh samples taken from a clinical trial, banked samples, or contrived samples that are representative of fresh samples. If this approach is used, we recommend that samples be split so replicate analysis by every Trained User and Untrained User who participates in the study can be further assessed. We also recommend that you include a discussion of rationale for use of contrived or banked samples, including human factors, if appropriate.

Suggestions for this approach are provided below.

- **Sites**. A minimum of three study sites should be included.
- **Samples**. There should be a set of samples that are within, and span, the measuring range for the test. Samples should also be analyzed under conditions that mimic the conditions under which patient-drawn samples would be performed (e.g., at different times during the work day, introduced among other tasks performed by Untrained Users). The CLIA waiver applicant should describe how it determined the number of samples, and the selection of samples, to use in the study, and the steps taken to mimic realistic testing conditions.
- **Untrained Users**. The number of Untrained Users should be sufficient to ensure a representative user population, and sufficient statistical power to allow for a fulsome evaluation of the tests to a statistical significance. Generally, a study with nine Untrained Users should be sufficient, though a sponsor may choose to use more.
- **Trained Users**. If the sponsor chooses not to use data from a prior 510(k) or PMA study, to establish Trained User performance, a minimum of six Trained Users should generally participate . However, a lesser number of users might be appropriate if results are not dependent on visual interpretation.
- **Analysis of Samples**. Each sample should be independently analyzed by every Trained User and Untrained User by site, so direct comparison of operator impact can be performed. For example, if a set of 40 samples is used, and the number of participating Untrained Users and Trained Users were nine and six, respectively, there would be 360 total Untrained User results (40 samples in the set * 9 Untrained Users) and 240 Trained User results (40 samples in the set * 6 Trained Users). In addition, batch testing of Trained User samples from multiple sites at a single central laboratory site may be considered.

- **Statistical Analysis**. Limits of agreement with the known values and their 95% confidence intervals should be derived for the Trained Users and the Untrained Users using a similar approach as in 3.1, adapted to reflect the fact that the true (assigned) value of each sample is known. A predetermined value of the width of the 95% confidence intervals should be used to assess if the limits of agreement for each group are within the acceptable range. For additional general guidance of statistical methods for quantitative tests, please refer to CLSI EP9.

4. Qualitative IVD Studies

Two options for CLIA waiver accuracy evaluations of qualitative tests are discussed in this section –

1. Agreement studies using subjects enrolled in a CLIA waiver clinical trial; and
2. Equivalence of accuracy studies using samples with assigned values.

The sponsor should determine which option is most appropriate for their IVD. Also, sponsors are encouraged to discuss alternative approaches, or hybrids of these approaches, with FDA that are appropriate for their particular technology and the IVD's intended use. For example, each of the options described below assume that both Trained Users and Untrained Users will participate in the study, though as discussed in Section 2.6, there may be approaches where data or criteria from 510(k)s or PMAs may be used in lieu of Trained Users in the CLIA waiver study. In those instances, sponsors may want to discuss adopting the study designs described below to leverage the pre-existing data or criteria.

4.1 Agreement Study with Subjects and Prepared Samples

An agreement study as described below might be used to evaluate a qualitative test's accuracy, though in general we recommend the following approach as an option, though others may be used. An agreement study can provide comprehensive picture of accuracy for a qualitative test, and is intended to isolate any effect of the user on the test, and to provide adequate assurance that a product is "accurate" for CLIA waiver purposes.

Suggestions for this study design are provided below.

- **Sites**. A minimum of three study sites should be included.
- **Samples**. There are two sample sets that should be used
    - **Subject Samples**. It is generally appropriate to perform a study using a sufficient number of positive and negative samples to provide a meaningful assessment of performance between Trained Users and Untrained Users. Generally, 120 positive and 120 negative samples total across all sites should be sufficient for analysis.

      Preferably, samples should be obtained from subjects *representing the real-world presentation of patients*. For example, if a test will be used as an aid in

diagnosis or to monitor patients with known conditions, the subjects in the study should have those attributes, as those who reflect the real world conditions of use. Another option may be to accept subjects who vary in some respects from common presentation of patients (e.g., subjects with a confirmed positive infection for a screening test) who could provide additional samples to span the range of analysis. To the extent that it is not possible to identify a sufficient number of positive or negative samples in subjects with reasonable efforts, banked or contrived samples may be used to supplement the sample set.

- **Prepared Samples**. In addition, the manufacturer should prepare an additional 60 weak positive and 60 negative weak samples of known concentration. The sponsor should prepare samples with the concentration range in which results change, e.g., positive to negative or reactive to unreactive. Assessment is best made at the concentrations at which, ideally, 5% and 95% (or in some instances, a more flexible range of 30% to 70%) of the measurements for that sample are either positive or negative, as described in CLSI EP-12-A2. However, we have recognized through our review of qualitative test data, that as methodologies have progressed and improved, the transition from negative to positive may be so "sharp" that preparation of samples that produce these suggested ranges might be impractical. Therefore, depending upon the analytical and clinical conditions of use, different concentrations that are still within the transition zone, or as a minimum, concentrations that clearly demonstrate performance across the transition range (e.g., there is a "very small" concentration interval between positive and negative or reactive and unreactive) might be acceptable..

- **Untrained Users**. The number of Untrained Users should be sufficient to ensure a representative user population, and sufficient statistical power to allow for a fulsome evaluation of the tests to a statistical significance. Generally, a study with nine Untrained Users should be sufficient, though a sponsor may choose to use more.
- **Trained Users**. If the sponsor chooses not to use data from a prior 510(k) or PMA study, to establish Trained User performance, a minimum of six Trained Users should generally participate . However, lesser number of users might be appropriate if results are not dependent on visual interpretation.
- **Analysis of Samples**. For each subject, samples should be independently collected and analyzed by any one Untrained User and any one Trained User at a study site. In addition, batch testing of Trained User samples from multiple sites at a single central laboratory site may be considered.
- **Statistical Analysis**: Analyses are based on an assessment of positive percent agreement (PPA) and negative percent agreement (NPA) between the Untrained User and Trained User groups. Trained User results and Untrained User results should be pooled within each site.

- **Subject Sample Data Set**. The observed PPA and NPA for Trained User and Untrained User populations at each site should generally achieve minimum agreement of 90%, and a lower confidence bound in the range of 80%, although greater or lesser agreement might be required depending upon the test; minimum agreement criteria should be determined taking Sections 2.6.2 and 2.6.3 into consideration. We also recommend that you report any invalid results obtained. This can be summarized using a table or graph that captures the number of times a sample and/or a specimen needed to be tested in order to obtain a valid result.
- **Prepared Sample Data Set**. Determine whether samples can be pooled across sites using the Fisher-Freeman-Halton test. If samples can be pooled, evaluate as follows. The percent of positive results for the 60 aliquots of the prepared, weak positive samples should generally be 85%, although greater or less agreement might be required depending upon the test. That is, approximately 51 out of 60 of these samples should yield positive results. The percent negative results for the prepared weak negative samples should also be 85%. Minimum agreement criteria should be determined taking Sections 2.6.2 and 2.6.3 into consideration.

  If results cannot be pooled, and each site individually does not achieve minimum agreement criteria, the sponsor should evaluate the reasons and determine how performance might be improved at underperforming site(s).

For additional general guidance of statistical methods for qualitative tests, please refer to CLSI EP12.

4.2 Accuracy studies using samples with assigned values

As noted in Section 3.2, it is possible to compare the accuracy of Trained Users and Untrained Users using a set of samples in which all users analyze each sample. A study design for qualitative tests using similar concepts is provided below.

- **Sites**. A minimum of three study sites should be included.
- **Samples**. There should be a set of samples including multiple weak positive and weak negative samples. Samples should also be analyzed under conditions that mimic the conditions under which patient-drawn samples would be performed (e.g., at different times during the work day, introduced among other tasks performed by Untrained Users). The CLIA waiver applicant should describe how it determined the number of samples, and the selection of samples, to use in the study and the steps taken to mimic realistic testing conditions.
- **Untrained Users**. The number of Untrained Users should be sufficient to ensure a representative user population, and sufficient statistical power to allow for a fulsome

evaluation of the tests to a statistical significance. Generally, a study with nine Untrained Users should be sufficient, though a sponsor may choose to use more.

- **Trained Users**. If the sponsor chooses not to use data from a prior 510(k) or PMA study, to establish Trained User performance, a minimum of six Trained Users should generally participate. However, a lesser number of users might be appropriate if results are not dependent on visual interpretation.
- **Analysis of Samples**. Each sample should be independently analyzed by every Trained User and Untrained User so comparison of operator population effects can be assessed for each sample. For example, if a set of 40 samples is used, and the number of participating Untrained Users and Trained Users were nine and six, respectively, there would be 360 total Untrained User results (40 samples in the set * 9 Untrained Users) and 240 Trained User results (40 samples in the set * 6 Trained Users). In addition, batch testing of Trained User samples from multiple sites at a single central laboratory site may be considered.
- **Statistical Analysis**. Sensitivity and specificity should be calculated for each Trained User and Untrained User, and pooled if appropriate. The equivalence between the Trained User and Untrained User user groups should be determined by considering whether the expected value of the squared difference between a Trained User and a Untrained User does not exceed the expected value of the squared difference between two Trained Users by a predetermined margin.[8] The margin should be determined taking sections 2.6.1 and 2.6.3 into account. For additional general guidance of statistical methods for qualitative tests, please refer to CLSI EP12.

5. Special Considerations

5.1. Samples with Assigned Values

As discussed above, sponsors may in some cases use samples with assigned values. The assignment would preferably be determined in one of two ways. If fresh patient samples are used, analyte concentrations would be determined using a recognized reference method or, alternatively by a recognized commercial (e.g., FDA cleared or approved) method. Another option would be to use contrived samples that would be prepared using known amounts of analyte.

To facilitate the use of assigned-value standards in these instances, FDA will work with sponsors to develop a study design that complies with least burdensome principles, as needed.

[8] *See* Schall R and Luus H, *On population and individual bioequivalence*. Statistics in Medicine 1993, pp1109-1124; Obuchowski N, *Can electronic medical images replace hard copy? Defining and testing the equivalence of diagnostic tests*. Statistics in Medicine 2001, pp 2845-2863.

5.2. Differences in User Performance

In some instances, there might be differences in the performance of Trained Users and Untrained Users that rise to a level of statistical significance, but is not clinically meaningful. There also may be instances where differences in performance exist, but the public health benefits or the low level of risk warrant allowing differences in performance at certificate of waiver sites. Sponsors may submit justifications, and FDA may waive tests in these instances, and FDA may determine that tests are sufficiently "accurate" (and "simple") to support a waiver under 42 USC § 263a(d)(3)(A), or be waived under other provisions.[9]

5.3. Dual Submissions

The principles with respect to determining accuracy apply equally to CLIA application submitted after initial clearance or approval of an IVD or to dual submissions seeking clearance and waivers at the same time. However, for dual submission, modifications may be made to clearance and approval studies to reduce the burden of analysis. We recommend discussing test design with FDA if the sponsor plans to take this approach.

[9] As written, the CLIA waiver law allows FDA to waive tests "*including* those that – (A) employ methodologies that are so simple and accurate as to render the likelihood of erroneous results by the user negligible, or (B) [FDA] has determined pose no unreasonable risk of harm to the patient if performed incorrectly." 42 USC § 263a(d)(3) (emphasis added). The term "including" indicates that Congress did not create just two pathways for waiver (pathway (A) or pathway (B)), but allowed for the possibility of other pathways as well.